# Introduction to Matrices

**Definition**

① The **trace** $\mathrm{tr}(A)$ of an $n \times n$ matrix $A = (a_{ij})$ is the sum of the diagonal entries of $A$. That is, $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

② The transpose $A^{\mathrm{T}}$ of an $m \times n$ matrix $A$ is the $n \times m$ matrix whose $(i, j)$-entry is $a_{ji}$.

③ A matrix $A$ is called **symmetric** if $A^{\mathrm{T}} = A$.

④ We say two matrices $A, B$ **commute** if $AB = BA$.

⑤ The entries $a_{ii}$ of a matrix $A = (A_{ij})$ are called **diagonal entries**.

⑥ A **diagonal matrix** is a square matrix whose non-diagonal entries are all zero.

⑦ An $n \times n$ matrix whose diagonal entries are all $1$ is called the **identity matrix** and denoted by $I_n$, or simply by $I$.

**Summary**

① For matrices $A, B$, and a scalar $r$, the followings are true when the expressions defined:

1. $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$
2. $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$
3. $(rA)^{\mathrm{T}} = rA^{\mathrm{T}}$

② For $n \times n$ matrices $A, B$, and a scalar $r$, we have

1. $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$
2. $\mathrm{tr}(rA) = r\mathrm{tr}(A)$

## Problems

① Let $A$ and $B$ are matrices such that the matrix product $AB$ is defined and $AB$ is a square matrix. Is it true that the matrix product $BA$ is also defined and $BA$ is a square matrix? If it is true, then prove it. If not, find a counterexample.

### Definition/Hint.

Let $A$ be an $m \times n$ matrix.

This means that the matrix $A$ has $m$ rows and $n$ columns.

Let $B$ be an $r \times s$ matrix.

Then the matrix product $AB$ is defined if $n = r$, that is, if the number of columns of $A$ is equal to the number of rows of $B$.

*Definition.* A matrix $C$ is called a **square matrix** if the size of $C$ is $n \times n$ for some positive integer $n$.
(The number of rows and the number of columns are the same.)

## Proof.

We prove that the matrix product $BA$ is defined and it is a square matrix.

Let $A$ be an $m \times n$ matrix and $B$ be an $r \times s$ matrix.

Since the matrix product $AB$ is defined, we must have $n = r$ and the size of $AB$ is $m \times s$.
Since $AB$ is a square matrix, we have $m = s$.

Thus the size of the matrix $B$ is $n \times m$.
From this, we see that the product $BA$ is defined and its size is $n \times n$, hence it is a square matrix.

---

2 Let $A$ and $B$ be $2 \times 2$ matrices. Prove or find a counterexample for the statement that $(A - B)(A + B) = A^2 - B^2$.

## Hint.

In general, matrix multiplication is not commutative: $AB$ and $BA$ might be different.

## Solution.

Let us calculate $(A - B)(A + B)$ as follows using the fact that the matrix product is distributive.

$$\begin{aligned}(A-B)(A+B) &= A(A+B) - B(A+B) \\ &= A^2 + AB - BA - B^2 \\ &= A^2 - B^2 + (AB - BA).\end{aligned}$$

Thus if $(A - B)(A + B) = A^2 - B^2$ then $AB - BA = O$, the zero matrix. Equivalently, $AB = BA$.

---

Note that matrix multiplication is not commutative, namely, $AB \neq BA$ in general.
Thus we can disprove the statement if we find matrices $A$ and $B$ such that $AB \neq BA$.

---

For example, let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Then we have

$$AB = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \text{ and } BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since $AB \neq BA$, we have $(A - B)(A + B) \neq A^2 - B^2$ for

$$A=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } B=\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Hence we found a counterexample for the statement.

3 Let $A$ and $B$ be $n \times n$ matrices. Suppose that the matrix product $AB=O$, where $O$ is the $n \times n$ zero matrix. Is it true that the matrix product with opposite order $BA$ is also the zero matrix? If so, give a proof. If not, give a counterexample.

## Solution.

The statement is in general not true. We give a counter example.

Consider the following 2 × 2 matrices.

$$A=\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then we compute

$$AB=\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}=O.$$

Thus the matrix product $AB$ is the 2 × 2 zero matrix $O$.

On the other hand, we compute

$$BA=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

Thus the matrix product $BA$ is not the zero matrix.

Therefore the statement is not true in general.

4 Find a nonzero $3 \times 3$ matrix $A$ such that $A^2 \neq O$ and $A^3=O$, where $O$ is the $3 \times 3$ zero matrix.

## Solution.

For example, let $A$ be the following 3 × 3 matrix.

$$A=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then $A$ is a nonzero matrix and we have

$$A^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq O.$$

The third power of $A$ is

$$A^3 = A^2 A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = O.$$

Thus, the nonzero matrix $A$ satisfies the required conditions $A^2 \neq O, A^3 = O$.

## Comment.

A square matrix $A$ is called **nilpotent** if there is a non-negative integer $k$ such that $A^k$ is the zero matrix.
The smallest such an integer $k$ is called **degree** or **index** of $A$.

The matrix $A$ in the solution above gives an example of a $3 \times 3$ nilpotent matrix of degree 3.

---

5 Let

$$A = \begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix} \text{ and } \mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Compute $A^{2017}\mathbf{u}$.
(*The Ohio State University*)

## Solution.

We first compute $A\mathbf{u}$. We have

$$A\mathbf{u} = \begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} = -\begin{bmatrix} 1 \\ 0 \end{bmatrix} = -\mathbf{u}.$$

Thus we have

$$A\mathbf{u} = -\mathbf{u}.$$

Then using this result repeatedly we have

$$A^2\mathbf{u} = A(A\mathbf{u}) = A(-\mathbf{u}) = -A\mathbf{u} = -(-\mathbf{u}) = \mathbf{u}.$$

Thus we have

$$A^2\mathbf{u} = \mathbf{u}.$$

Next, we compute

$$A^3\mathbf{u} = A(A^2\mathbf{u}) = A(\mathbf{u}) = -\mathbf{u}.$$

Now you see the pattern and obtain

$$A^n\mathbf{u} = \begin{cases} -\mathbf{u} & \text{if } n \text{ is odd} \\ \mathbf{u} & \text{if } n \text{ is even.} \end{cases}$$

(To prove this pattern, use mathematical induction. See below.)

Thus we conclude that

$$A^{2017}\mathbf{u} = -\mathbf{u} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

since 2017 is odd.

### The proof of the formula.

For completeness, we prove the pattern (*) we see above.

We use mathematical induction.

When $n = 1$, we computed $A\mathbf{u} = -\mathbf{u}$ and the base case is true.

Suppose that(*) is true for $n = k$.

Then we have

$$\begin{aligned} A^{k+1}\mathbf{u} = A(A^k\mathbf{u}) &= \begin{cases} A(-\mathbf{u}) & \text{if } k \text{ is odd} \\ A(\mathbf{u}) & \text{if } k \text{ is even} \end{cases} \\ &= \begin{cases} -A\mathbf{u} & \text{if } k \text{ is odd} \\ A\mathbf{u} & \text{if } k \text{ is even} \end{cases} \\ &= \begin{cases} \mathbf{u} & \text{if } k \text{ is odd} \\ -\mathbf{u} & \text{if } k \text{ is even} \end{cases} \quad = \begin{cases} \mathbf{u} & \text{if } k+1 \text{ is even} \\ -\mathbf{u} & \text{if } k+1 \text{ is odd} \end{cases} \end{aligned}$$

Here the second equality follows from the induction hypothesis.

This proves (*) for $n = k + 1$ and this completes the induction step and (*) is true for any positive integer $n$.

---

6 Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

be a $3 \times 3$ matrix. Then find the formula for $A^n$ for any positive integer $n$.

### Proof.

We first compute several powers of $A$ and guess the general formula.

We have

$$A^2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A^3 = A^2A = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A^4 = A^3A = \begin{bmatrix} 1 & 1 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 7 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

From these computations, we guess the general formula of $A^n$ is

$$A^n = \begin{bmatrix} 1 & 1 & 2n-1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

---

We prove this formula by mathematical induction on $n$.

The base case $n = 1$ follows from the definition of $A$.

Suppose that the formula is true for $n = k$.

We prove the formula for $n = k + 1$.

We have

$$\begin{aligned} A^{k+1} &= A^kA \\ &= \begin{bmatrix} 1 & 1 & 2k-1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} && \text{by the induction hypothesis} \\ &= \begin{bmatrix} 1 & 1 & 2k+1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 2(k+1)-1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Thus the formula holds for $n = k + 1$.

Hence the formula is true for any positive integer $n$ by induction.

---

7 A square matrix $A$ is called **idempotent** if $A^2 = A$.

**(a)** Suppose $A$ is an $n \times n$ idempotent matrix and let $I$ be the $n \times n$ identity matrix. Prove that the matrix $I - A$ is an idempotent matrix.

**(b)** Assume that $A$ is an $n \times n$ nonzero idempotent matrix. Then determine all integers $k$ such that the matrix $I - kA$ is idempotent.

**(c)** Let $A$ and $B$ be $n \times n$ matrices satisfying $AB = A$ and $BA = B$. Then prove that $A$ is an idempotent matrix.

## Proof.

### (a) Prove that the matrix $I - A$ is an idempotent matrix.

To prove that $I - A$ is an idempotent matrix, we show that $(I - A)^2 = I - A$.

We compute

$$\begin{aligned}(I-A)^2 &= (I-A)(I-A)\\ &= I(I-A) - A(I-A)\\ &= I - A - A + A^2\\ &= I - 2A + A^2\\ &= I - 2A + A && \text{since } A \text{ is idempotent}\\ &= I - A.\end{aligned}$$

Thus, we have $(I - A)^2 = I - A$ and $I - A$ is an idempotent matrix.

### (b) Determine all integers $k$ such that the matrix $I - kA$ is idempotent.

Let us find out the condition on $k$ so that $I - kA$ is an idempotent matrix.

We have

$$\begin{aligned}&(I-kA)^2 = (I-kA)(I-kA)\\ &= I(I-kA) - kA(I-kA)\\ &= I - kA - kA + k^2A^2\\ &= I - 2kA + k^2A && \text{since } A \text{ is idempotent}\\ &= I - (2k - k^2)A.\end{aligned}$$

It follows that $I - kA$ is idempotent if and only if $I - kA = I - (2k - k^2)A$, or equivalently $(k^2 - k)A = O$, the zero matrix.

Since $A$ is not the zero matrix, we see that $I - kI$ is idempotent if and only if $k^2 - k = 0$.

Since $k^2 - k = k(k-1)$, we conclude that $I - kA$ is an idempotent matrix if and only if $k = 0, 1$.

### (c) Prove that $A$ is an idempotent matrix.

Let $A$ and $B$ be $n \times n$ matrices satisfying

$$AB = A \tag{*}$$

and

$$BA = B. \tag{**}$$

Our goal is to show that $A^2 = A$.

We compute

$$\begin{aligned}&A^2 = AA\\ &= (AB)A && \text{by (*)}\\ &= A(BA)\\ &= AB && \text{by (**)}\\ &= A && \text{by (*)}.\end{aligned}$$

Therefore, we have obtained the identity $A^2 = A$, and we conclude that $A$ is an idempotent matrix.

---

⑧ Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.

**(a)** Find all matrices $B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ such that $AB = BA$.

**(b)** Use the results of part (a) to exhibit $2 \times 2$ matrices $B$ and $C$ such that $AB = BA$ and $AC \neq CA$.

## Solution.

### Find all matrices $B$ such that $AB = BA$

Since we want to find $B$ such that $AB = BA$, we need to find $x, y, z, w$ satisfying

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}.$$

Comparing entries, we have the following system of linear equations.

$$\begin{aligned} x + 3z &= x + 2y \\ y + 3w &= 3x + 4y \\ 2x + 4z &= z + 2w \\ 2y + 4w &= 3z + 4w. \end{aligned}$$

Simplifying this, we obtain

$$\begin{aligned} 2y - 3z &= 0 \\ 3x + 3y - 3w &= 0 \\ 2x + 3z - 2w &= 0 \\ 2y - 3z &= 0. \end{aligned}$$

Note that the first and the last equations are identical. So we omit the first equation from our consideration.

To solve this system, we use the Gauss-Jordan elimination method. Namely we form the augmented matrix of this system and apply elementary row operations as follows.

$$\left[\begin{array}{cccc|c} 3 & 3 & 0 & -3 & 0 \\ 2 & 0 & 3 & -2 & 0 \\ 0 & 2 & -3 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 2 & 0 & 3 & -2 & 0 \\ 0 & 2 & -3 & 0 & 0 \end{array}\right] \xrightarrow{R_2 - 2R_1}$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & -2 & 3 & 0 & 0 \\ 0 & 2 & -3 & 0 & 0 \end{array}\right] \xrightarrow[R_2 + R_3]{R_1 + \frac{1}{2}R_2} \left[\begin{array}{cccc|c} 1 & 0 & 3/2 & -1 & 0 \\ 0 & -2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{-\frac{1}{2}R_2} \left[\begin{array}{cccc|c} 1 & 0 & 3/2 & -1 & 0 \\ 0 & 1 & -3/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The last matrix is in reduced row echelon form. From this, we obtain the general solution

$$
\begin{aligned}
x &= -\frac{3}{2}z + w \\
y &= \frac{3}{2}z,
\end{aligned}
$$

where $z$ and $w$ are free variables.

Therefore the matrix $B$ such that $AB = BA$ must be

$$
\begin{aligned}
B &= \begin{bmatrix} -\frac{3}{2}z + w & \frac{3}{2}z \\ z & w \end{bmatrix} \\
&= z \begin{bmatrix} -\frac{3}{2} & \frac{3}{2} \\ 1 & 0 \end{bmatrix} + w \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

for any $z$ and $w$.

### (b) Find matrices $B, C$ such that $AB = BA$, $AC \neq CA$

By part (a), if

$$B = z \begin{bmatrix} -\frac{3}{2} & \frac{3}{2} \\ 1 & 0 \end{bmatrix} + w \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

for some $z, w$, then we have $AB = BA$.

For example, let $z = 2, w = 0$. Then the matrix

$$B = \begin{bmatrix} -3 & 3 \\ 2 & 0 \end{bmatrix}$$

satisfies $AB = BA$.

To find a matrix $C$ such that $AC \neq CA$, the matrix $C$ must not be of the form of the formula of $B$.

For example, let

$$C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

You may directly check that $AC \neq CA$.

Or, we can show that $C$ is never be the matrix of the form

$$\begin{bmatrix} -\frac{3}{2}z + w & \frac{3}{2}z \\ z & w \end{bmatrix}.$$

To see this, compare $(1, 2)$ and $(2, 1)$ entries. There is no $z$ such that

$$\frac{3}{2}z = 0 \text{ and } z = 1.$$

(This is how I found the matrix $C$ as above.)

⑨ Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 \\ 0 & 6 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Then compute and simplify the following expression.

$$\mathbf{v}^{\mathrm{T}} \left( A^{\mathrm{T}} - (A - B)^{\mathrm{T}} \right) C.$$

**Solution.**

### (b) Matrix product and transpose

First of all, note that by the property of the transpose we have

$$(A - B)^{\mathrm{T}} = A^{\mathrm{T}} - B^{\mathrm{T}}.$$

Hence we can simply the middle part:

$$\begin{aligned} A^{\mathrm{T}} - (A - B)^{\mathrm{T}} &= A^{\mathrm{T}} - (A^{\mathrm{T}} - B^{\mathrm{T}}) \\ &= A^{\mathrm{T}} - A^{\mathrm{T}} + B^{\mathrm{T}} = B^{\mathrm{T}}. \end{aligned}$$

Thus the expression becomes

$$\mathbf{v}^{\mathrm{T}} \left( A^{\mathrm{T}} - (A - B)^{\mathrm{T}} \right) C = \mathbf{v}^{\mathrm{T}} B^{\mathrm{T}} C.$$

The transposes of $\mathbf{v}$ and $B$ are

$$\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \text{ and } B^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Thus we have

$$\begin{aligned} \mathbf{v}^{\mathrm{T}} \left( A^{\mathrm{T}} - (A - B)^{\mathrm{T}} \right) C &= \mathbf{v}^{\mathrm{T}} B^{\mathrm{T}} C \\ &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 6 \end{bmatrix}. \end{aligned}$$

In conclusion, we have obtained

$$\mathbf{v}^{\mathrm{T}} \left( A^{\mathrm{T}} - (A - B)^{\mathrm{T}} \right) C = \begin{bmatrix} 0 & 6 \end{bmatrix}.$$

---

⑩ Let $\mathbf{v} = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix}$. Find all $3 \times 1$ column vectors $\mathbf{w}$ such that $\mathbf{v}\mathbf{w} = 0$.

**Solution.**

Let $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$.

Then we want

$$\mathbf{v}\mathbf{w} = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = 2w_1 - 5w_2 - w_3 = 0.$$

Letting $w_1$ and $w_2$ be free variables, we solve $w_3 = 2w_1 - 5w_2$.

Then every solution to the equation $\mathbf{v}\mathbf{w} = 0$ is of the form

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ 2w_1 - 5w_2 \end{bmatrix} = w_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + w_2 \begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix}.$$

---

⑪ Prove the following identity for any positive integer $n$.

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^n = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}.$$

### Hint.

Recall the addition theorem of trigonometric functions (sum formula)

$$\begin{aligned} \sin(x+y) &= \sin x \cos y + \cos x \sin y \\ \cos(x+y) &= \cos x \cos y - \sin x \sin y. \end{aligned}$$

Use the induction on $n$.

### Proof.

We prove the identity by induction on $n$.

The base case $n = 1$ is clear.

Suppose that the identity is true for $n = k$.

Then we have

$$\begin{aligned} &\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{k+1} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^k \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{bmatrix} \\ &\text{(by the induction hypothesis)} \\ &= \begin{bmatrix} \cos\theta\cos k\theta - \sin\theta\sin k\theta & -\cos\theta\sin k\theta - \sin\theta\cos k\theta \\ \sin\theta\cos k\theta + \cos\theta\sin k\theta & -\sin\theta\sin k\theta + \cos\theta\cos k\theta \end{bmatrix} \\ &\text{(by matrix multiplication)} \\ &= \begin{bmatrix} \cos(\theta + k\theta) & -\sin(\theta + k\theta) \\ \sin(\theta + k\theta) & \cos(\theta + k\theta) \end{bmatrix} \\ &\text{(by the addition theorem of trigonometric functions)} \\ &= \begin{bmatrix} \cos(k+1)\theta & -\sin(k+1)\theta \\ \sin(k+1)\theta & \cos(k+1)\theta \end{bmatrix}. \end{aligned}$$

This proves that the identity holds for $n = k + 1$.

Thus by induction, the identity is true for all positive integers $n$.

---

⑫ Let $A$ and $B$ be $n \times n$ matrices. Is it always true that $\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B)$? If it is true, prove it. If not, give a counterexample.

**Solution.**

There are many counterexamples.

For one, take

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then $\operatorname{tr}(A) = 1, \operatorname{tr}(B) = 1$, and hence $\operatorname{tr}(A)\operatorname{tr}(B) = 1$, while $\operatorname{tr}(AB) = 0$ as $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

---

⑬ Let $A$ be an $n \times n$ matrix. Is it true that $\operatorname{tr}(A^{\mathrm{T}}) = \operatorname{tr}(A)$? If it is true, prove it. If not, give a counterexample.

**Solution.**

The answer is true. Recall that the transpose of a matrix is the sum of its diagonal entries. Also, note that the diagonal entries of the transposed matrix are the same as the original matrix.

Putting together these observations yields the equality $\operatorname{tr}(A^{\mathrm{T}}) = \operatorname{tr}(A)$.

---

Here is the more formal proof.

For $A = (a_{ij})_{1 \le i,j \le n}$, the transpose $A^{\mathrm{T}} = (b_{ij})_{1 \le i,j \le n}$ is defined by $b_{ij} = a_{ji}$.

In particular, notice that $b_{ii} = a_{ii}$ for $1 \le i \le n$. And so,

$$\operatorname{tr}(A^{\mathrm{T}}) = \sum_{i=1}^{n} b_{ii} = \sum_{i=1}^{n} a_{ii} = \operatorname{tr}(A).$$

---

⑭ Let $\mathbf{v}$ and $\mathbf{w}$ be two $n \times 1$ column vectors.
**(a)** Prove that $\mathbf{v}^{\mathrm{T}}\mathbf{w} = \mathbf{w}^{\mathrm{T}}\mathbf{v}$.
**(b)** Provide an example to show that $\mathbf{v}\mathbf{w}^{\mathrm{T}}$ is not always equal to $\mathbf{w}\mathbf{v}^{\mathrm{T}}$.

**Solution.**

(a) Prove that $\mathbf{v}^{\mathrm{T}}\mathbf{w} = \mathbf{w}^{\mathrm{T}}\mathbf{v}$.

Suppose the vectors have component

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

Then,

$$\mathbf{v}^{\mathrm{T}}\mathbf{w} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \sum_{i=1}^{n} v_i w_i,$$

while

$$\mathbf{w}^{\mathrm{T}}\mathbf{v} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} w_i v_i.$$

We can see that they are equal because $v_i w_i = w_i v_i$.

(b) Provide an example to show that $\mathbf{v}\mathbf{w}^{\mathrm{T}}$ is not always equal to $\mathbf{w}\mathbf{v}^{\mathrm{T}}$.

For the counterexample, let $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then

$$\mathbf{v}\mathbf{w}^{\mathrm{T}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

while

$$\mathbf{w}\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

### Comment.

Recall that for two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, the **dot product** (or **inner product**) of $\mathbf{v}, \mathbf{w}$ is defined to be

$$\mathbf{v} \cdot \mathbf{w} := \mathbf{v}^{\mathrm{T}}\mathbf{w}.$$

Part (a) of the problem deduces that the dot product is commutative. This means that we have

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}.$$

In fact, we have

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^{\mathrm{T}}\mathbf{w} \overset{(a)}{=} \mathbf{w}^{\mathrm{T}}\mathbf{v}\mathbf{w} \cdot \mathbf{v}.$$

Also, notice that while $\mathbf{v}\mathbf{w}^{\mathrm{T}}$ is not always equal to $\mathbf{w}\mathbf{v}^{\mathrm{T}}$, we know that $(\mathbf{v}\mathbf{w}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{w}\mathbf{v}^{\mathrm{T}}$.

**15** Let $\mathbf{v}$ and $\mathbf{w}$ be two $n \times 1$ column vectors. Prove that $\mathrm{tr}(\mathbf{v}\mathbf{w}^{\mathrm{T}}) = \mathbf{v}^{\mathrm{T}}\mathbf{w}$.

### Solution.

Suppose the vectors have components

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

Then,

$$\begin{aligned} \mathbf{v}\mathbf{w}^{\mathrm{T}} &= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \\ &= \begin{bmatrix} v_1 w_1 & v_1 w_2 & \cdots & v_1 w_n \\ v_2 w_1 & v_2 w_2 & \cdots & v_2 w_n \\ \vdots & \vdots & \vdots & \vdots \\ v_n w_1 & v_n w_2 & \cdots & v_n w_n \end{bmatrix}. \end{aligned}$$

We can now see that

$$\operatorname{tr}(\mathbf{v}\mathbf{w}^{\mathrm{T}}) = \sum_{i=1}^{n} v_i w_i = \mathbf{v}^{\mathrm{T}}\mathbf{w}.$$

### Comment.

Recall that $\mathbf{v}^{\mathrm{T}}\mathbf{w}$ is, by definition, the dot product of the vectors $\mathbf{v}$ and $\mathbf{w}$.

So, the dot product of vectors $\mathbf{v}$ and $\mathbf{w}$ is equal to the trace of the matrix $\mathbf{v}\mathbf{w}^{\mathrm{T}}$.

---

16 Let $\mathbf{v}$ be an $n \times 1$ column vector. Prove that $\mathbf{v}^{\mathrm{T}}\mathbf{v} = 0$ if and only if $\mathbf{v}$ is the zero vector $\mathbf{0}$.

### Proof.

Let $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

Then we have

$$\mathbf{v}^{\mathrm{T}}\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} \mathbf{v}_i^2.$$

Because each $v_i^2$ is non-negative, this sum is $0$ if and only if $v_i = 0$ for each $i$. In this case, $\mathbf{v}$ is the zero vector.

### Comment.

Recall that the the **length** of the vector $\mathbf{v} \in \mathbb{R}^n$ is defined to be

$$\|\mathbf{v}\| := \sqrt{\mathbf{v}^{\mathrm{T}}\mathbf{v}}.$$

The problem implies that the length of a vector is $0$ if and only if the vector is the zero vector.

---

17 Is it true that a real square matrix $A$ must commute with its transpose $A^{\mathrm{T}}$?

### Solution.

(a) Is it true that $A$ must commute with its transpose?

The answer is no.

We give a counterexample. Let

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

Then the transpose of $A$ is

$$A^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}.$$

We compute

$$AA^{\mathrm{T}} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix},$$

and

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}.$$

Therefore, we see that

$$AA^{\mathrm{T}} \neq A^{\mathrm{T}}A,$$

that is, $A$ does not commute with its transpose $A^{\mathrm{T}}$.

---

18 Let $\mathbf{v}$ be an $n \times 1$ column vector. Prove that $\mathbf{v}\mathbf{v}^{\mathrm{T}}$ is a symmetric matrix.

### Definition (Symmetric Matrix).

A matrix $A$ is called **symmetric** if $A^{\mathrm{T}} = A$.

In terms of entries, an $n \times n$ matrix $A = (a_{ij})$ is symmetric if $a_{ij} = a_{ji}$ for all $1 \leq i, j \leq n$.

### Proof.

Let $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. Then we have

$$\mathbf{v}\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 v_1 & v_1 v_2 & \cdots & v_1 v_n \\ v_2 v_1 & v_2 v_2 & \cdots & v_2 v_n \\ \vdots & \vdots & \vdots & \vdots \\ v_n v_1 & v_n v_2 & \cdots & v_n v_n \end{bmatrix}.$$

In particular, the the $(i, j)$-th component is

$$(\mathbf{v}\mathbf{v}^{\mathrm{T}})_{ij} = v_i v_j = v_j v_i = (\mathbf{v}\mathbf{v}^{\mathrm{T}})_{ji}.$$

This shows that the matrix $\mathbf{v}\mathbf{v}^{\mathrm{T}}$ is symmetric.

---

⑲ Calculate the following expressions, using the following matrices:

$$A = \begin{bmatrix} 2 & 3 \\ -5 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$$

**(a)** $AB^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}$.
**(b)** $A\mathbf{v} - 2\mathbf{v}$.
**(c)** $\mathbf{v}^{\mathrm{T}}B$.
**(d)** $\mathbf{v}^{\mathrm{T}}\mathbf{v} + \mathbf{v}^{\mathrm{T}}BA^{\mathrm{T}}\mathbf{v}$.

### Solution.

### (a) $AB^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}$

$$\begin{aligned} AB^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}} &= \begin{bmatrix} 2 & 3 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -4 \end{bmatrix} \begin{bmatrix} 2 & -4 \end{bmatrix} \\ &= \begin{bmatrix} -3 & -1 \\ -1 & -6 \end{bmatrix} + \begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -9 \\ -9 & 10 \end{bmatrix} \end{aligned}$$

### (b) $A\mathbf{v} - 2\mathbf{v}$

$$\begin{aligned} A\mathbf{v} - 2\mathbf{v} &= \begin{bmatrix} 2 & 3 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} -8 \\ -14 \end{bmatrix} + \begin{bmatrix} -4 \\ 8 \end{bmatrix} \\ &= \begin{bmatrix} -12 \\ -6 \end{bmatrix} \end{aligned}$$

### (c) $\mathbf{v}^{\mathrm{T}}B$

$$\begin{aligned}\mathbf{v}^{\mathrm{T}}B &= \begin{bmatrix} 2 & -4 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 2 \end{bmatrix}\end{aligned}$$

(d) $\mathbf{v}^{\mathrm{T}}\mathbf{v} + \mathbf{v}^{\mathrm{T}}BA^{\mathrm{T}}\mathbf{v}$

$$\begin{aligned}&\mathbf{v}^{\mathrm{T}}\mathbf{v} + \mathbf{v}^{\mathrm{T}}BA^{\mathrm{T}}\mathbf{v} \\ &= \begin{bmatrix} 2 & -4 \end{bmatrix}\begin{bmatrix} 2 \\ -4 \end{bmatrix} + \left(\begin{bmatrix} 2 & -4 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}\right)\left(\begin{bmatrix} 2 & -5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ -4 \end{bmatrix}\right) \\ &= 18 + \begin{bmatrix} -4 & 2 \end{bmatrix}\begin{bmatrix} 24 \\ 2 \end{bmatrix} \\ &= 18 - 92 \\ &= -74\end{aligned}$$

**20** Let $A$ and $B$ be $n \times n$ matrices, and $\mathbf{v}$ an $n \times 1$ column vector. Use the matrix components to prove that $(A+B)\mathbf{v} = A\mathbf{v} + B\mathbf{v}$.

**Solution.**

We will use the matrix components $A = (a_{ij})_{1 \le i,j \le n}$, $B = (b_{ij})_{1 \le i,j \le n}$, and $\mathbf{v} = (v_i)_{1 \le i \le n}$.

Then

$$(A+B)\mathbf{v} = \begin{bmatrix} \sum_{j=1}^{n}(a_{1j}+b_{1j})v_j \\ \sum_{j=1}^{n}(a_{2j}+b_{2j})v_j \\ \sum_{j=1}^{n}(a_{3j}+b_{3j})v_j \\ \vdots \\ \sum_{j=1}^{n}(a_{nj}+b_{nj})v_j \end{bmatrix}.$$

On the other hand,

$$A\mathbf{v} + B\mathbf{v} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}v_j \\ \sum_{j=1}^{n} a_{2j}v_j \\ \sum_{j=1}^{n} a_{3j}v_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}v_j \end{bmatrix} + \begin{bmatrix} \sum_{j=1}^{n} b_{1j}v_j \\ \sum_{j=1}^{n} b_{2j}v_j \\ \sum_{j=1}^{n} b_{3j}v_j \\ \vdots \\ \sum_{j=1}^{n} b_{nj}v_j \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n}(a_{1j}+b_{1j})v_j \\ \sum_{j=1}^{n}(a_{2j}+b_{2j})v_j \\ \sum_{j=1}^{n}(a_{3j}+b_{3j})v_j \\ \vdots \\ \sum_{j=1}^{n}(a_{nj}+b_{nj})v_j \end{bmatrix}.$$

We can see that they are the same.

**21** Let $A$ be a square matrix such that $A^{\mathrm{T}}A = A$, where $A^{\mathrm{T}}$ is the transpose matrix of $A$. Prove that $A$ is idempotent, that is, $A^2 = A$. Also, prove that $A$ is a symmetric matrix.

**Hint.**

Recall the basic properties of transpose matrices.

For matrices $A, B$, we have

㉒ $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$.

㉓ $A^{\mathrm{TT}} = A$.

### Proof.

We first prove that $A$ is a symmetric matrix.

We have

$$\begin{aligned} A^{\mathrm{T}} &= (A^{\mathrm{T}} A)^{\mathrm{T}} \\ &= A^{\mathrm{T}} A^{\mathrm{TT}} && \text{by property 1} \\ &= A^{\mathrm{T}} A && \text{by property 2} \\ &= A. \end{aligned}$$

Hence we obtained $A^{\mathrm{T}} = A$, and thus $A$ is a symmetric matrix.

---

Now we prove that $A$ is idempotent.

We compute

$$\begin{aligned} A^2 &= AA \\ &= A^{\mathrm{T}} A && \text{since } A \text{ is symmetric} \\ &= A && \text{by assumption.} \end{aligned}$$

Therefore, the matrix $A$ satisfies $A^2 = A$, and hence it is idempotent.

㉔ Let $A$ and $B$ be $n \times n$ real symmetric matrices. Prove the followings.
**(a)** The product $AB$ is symmetric if and only if $AB = BA$.
**(b)** If the product $AB$ is a diagonal matrix, then $AB = BA$.

### Hint.

A matrix $A$ is called ***symmetric*** if $A = A^{\mathrm{T}}$.

In this problem, we need the following property of transpose:

Let $A$ be an $m \times n$ and $B$ be an $n \times r$ matrix. Then we have

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}.$$

(When you distribute transpose over the product of two matrices, then you need to reverse the order of the matrix product.)

### Solution.

(a) The product $AB$ is symmetric if and only if $AB = BA$.

Suppose $AB$ is symmetric. This means that we have

$$AB = (AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}} = BA.$$

The second equality is a general property of transpose. In the last equality we used the assumption that $A, B$ are symmetric, that is, $A = A^{\mathrm{T}}, B = B^{\mathrm{T}}$. Thus we have $AB = BA$.

On the other hand, if $AB = BA$, then we have

$$AB = BA = B^{\mathrm{T}}A^{\mathrm{T}} = (AB)^{\mathrm{T}}.$$

Thus, we have $AB = (AB)^{\mathrm{T}}$, hence $AB$ is symmetric.

If the product $AB$ is a diagonal matrix, then $AB = BA$.

Note that a diagonal matrix is symmetric. Hence the result follows from part (a).

---

25 A $2 \times 2$ matrix has two parallel columns and $\mathrm{tr}(A) = 5$. Find $\mathrm{tr}(A^2)$.

### Solution.

For (a) and (b), we give two solutions. The first one does not use the knowledge of eigenvalues, and the second one uses eigenvalues.

Solution 1 of (a) A $2 \times 2$ matrix $A$ satisfies $\mathrm{tr}(A^2) = 5$ and $\mathrm{tr}(A) = 3$. Find $\det(A)$.

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then we have

$$A^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{bmatrix}.$$

Since $\mathrm{tr}(A^2) = 5$ and $\mathrm{tr}(A) = 3$, we obtain

$$5 = \mathrm{tr}(A^2) = (a^2 + bc) + (bc + d^2) = a^2 + 2bc + d^2 \text{ and}$$
$$3 = \mathrm{tr}(A) = a + d.$$

We find the determinant $\det(A) = ad - bc$ as follows.

We have

$$\det(A) = ad - bc = \frac{1}{2}\left((a + d)^2 - (a^2 + 2bc + d^2)\right)$$
$$= \frac{1}{2}(3^2 - 5) = 2.$$

Thus, we obtain $\det(A) = 2$.

### Solution 2 of (a)

Let $\lambda_1$ and $\lambda_2$ be eigenvalues of $A$.
Then we have

$$\begin{aligned} 3 &= \mathrm{tr}(A) = \lambda_1 + \lambda_2 \text{ and} \\ 5 &= \mathrm{tr}(A^2) = \lambda_1^2 + \lambda_2^2. \end{aligned}$$

Here we used two facts.
The first one is that the trace of a matrix is the sum of all eigenvalues of the matrix.
The second one is that $\lambda^2$ is an eigenvalue of $A^2$ if $\lambda$ is an eigenvalue of $A$, and these are all the eigenvalues of $A^2$.

Since the determinant of $A$ is the product of eigenvalues of $A$, we have

$$\begin{aligned} \det(A) &= \lambda_1\lambda_2 \\ &= \frac{1}{2}\left((\lambda_1+\lambda_2)^2 - (\lambda_1^2+\lambda_2^2)\right) \\ &= \frac{1}{2}(3^2-5) \\ &= 2. \end{aligned}$$

Hence we have $\det(A) = 2$.

### Solution 1 of (b) A $2 \times 2$ matrix has two parallel columns and $\mathrm{tr}(A) = 5$. Find $\mathrm{tr}(A^2)$.

Since two columns are parallel, we can write $A$ as

$$A = \begin{bmatrix} a & ra \\ c & rc \end{bmatrix}.$$

Then we have

$$5 = \mathrm{tr}(A) = a + rc.$$

We use the formula in Solution 1 of (a) for $\mathrm{tr}(A^2)$ with $b = ra$ and $d = rc$, and we compute

$$\begin{aligned} \mathrm{tr}(A^2) &= a^2 + 2(ra) + (rc)^2 \\ &= (a+rc)^2 \\ &= 5^2 = 25. \end{aligned}$$

Thus, we find $\mathrm{tr}(A^2) = 25$.

### Solution 2 of (b)

Since two columns are parallel, the matrix $A$ is singular. Hence $A$ has an eigenvalue 0.

Since the sum of all the eigenvalues is $\mathrm{tr}(A) = 5$, we see that 0 and 5 are eigenvalues of $A$.

It follows that 0 and 25 are eigenvalues of $A^2$. Hence

$$\mathrm{tr}(A^2) = 0 + 25 = 25.$$

### Solution of (c) A $2 \times 2$ matrix $A$ has $\det(A) = 5$ and positive integer eigenvalues. What is the trace of $A$?

The product of eigenvalues of $A$ is the determinant $\det(A) = 5$.

Since eigenvalues are positive integers, it follows that 1 and 5 are eigenvalues of $A$.

It follows that

$$\mathrm{tr}(A) = 1 + 5 = 6.$$

---

26 Let $I$ be the $n \times n$ identity matrix, where $n$ is a positive integer. Prove that there are no $n \times n$ matrices $X$ and $Y$ such that $XY - YX = I$.

### Hint.

Suppose that such matrices exist and consider the trace of the matrix $XY - YX$.

Recall that the trace of a square matrix $A$ is the sum of the diagonal entries of $A$.

### Proof.

We use the following basic properties of the trace of matrices.

Let $A, B$ be $n \times n$ matrices and let $c$ be a scalar.

27 $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

28 $\mathrm{tr}(cA) = c\,\mathrm{tr}(A)$.

29 $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Seeking a contradiction, we assume that there are matrices $X$ and $Y$ such that $XY - YX = I$.

Then we take the trace of both sides and obtain

$$\begin{aligned} n &= \mathrm{tr}(I) \\ &= \mathrm{tr}(XY - YX) \\ &= \mathrm{tr}(XY) - \mathrm{tr}(YX) && \text{(by property (1), (2) of the trace)} \\ &= \mathrm{tr}(XY) - \mathrm{tr}(YX) && \text{(by property (3) of the trace)} \\ &= 0. \end{aligned}$$

Since $n$ is a positive integer, this is a contradiction.

Therefore, such matrices $X, Y$ do not exist.

---

30 Let $A=(a_{ij})$ and $B=(b_{ij})$ be $n\times n$ real matrices for some $n\in\mathbb{N}$.
**(a)** Express $\mathrm{tr}(AB^{\mathrm{T}})$ in terms of the entries of the matrices $A$ and $B$. Here $B^{\mathrm{T}}$ is the transpose matrix of $B$.
**(b)** Show that $\mathrm{tr}(AA^{\mathrm{T}})$ is the sum of the square of the entries of $A$.
**(c)** Show that if $A$ is nonzero symmetric matrix, then $\mathrm{tr}(A^2)>0$.

### Hint.

Review

31 the definition of the transpose of a matrix

32 the definition of matrix multiplication

33 the definition of a symmetric matrix

Then the proofs of these statement is straightforward computations.

### Proof.

(a) Express $\mathrm{tr}(AB^{\mathrm{T}})$ in terms of the entries of $A$ and $B$.

Here we use the following notation for an entry of a matrix: the $(i,j)$-entry of a matrix $C$ is denoted by $(C)_{i,j}$.

Then the $(i,j)$-entry of $AB^{\mathrm{T}}$ is $(AB^{\mathrm{T}})_{ij}=\sum_{k=1}^{n}a_{ik}b_{jk}$.

Thus we have

$$\mathrm{tr}(AB^{\mathrm{T}})=\sum_{l=1}^{n}(AB^{\mathrm{T}})_{ll}=\sum_{l=1}^{n}\sum_{k=1}^{n}a_{lk}b_{lk}.$$

(b) Show that $\mathrm{tr}(AA^{\mathrm{T}})$ is the sum of the square of the entries of $A$

By the formula obtained in part (a), we have

$$\mathrm{tr}(AA^{\mathrm{T}})=\sum_{l=1}^{n}\sum_{k=1}^{n}a_{lk}^2.$$

This is the sum of the squares of entries of $A$.

(c) Show that if $A$ is nonzero symmetric matrix, then $\mathrm{tr}(A^2)>0$.

Since $A$ is a symmetric matrix, we have $A^{\mathrm{T}}=A$.

Thus by the result of part (b), we have

$$\operatorname{tr}(A^2)=\operatorname{tr}(AA^{\mathrm{T}})=\sum_{l=1}^{n}\sum_{k=1}^{n}a_{lk}^2>0.$$

The last sum is strictly positive since $A$ is not the zero matrix, there is a nonzero entry of $A$ (and of course the square of a real number is nonnegative).

## Comment.

The results we proved in this article can be extended to complex matrices, matrices with complex number entries.
In this case, the condition in (c) that $A$ is symmetric is replaced by the condition that $A$ is a hermitian matrix.

---

**34** For each of the following matrix $A$, prove that $\mathbf{x}^{\mathrm{T}}A\mathbf{x}\geq 0$ for all vectors $\mathbf{x}$ in $\mathbb{R}^2$. Also, determine those vectors $\mathbf{x}\in\mathbb{R}^2$ such that $\mathbf{x}^{\mathrm{T}}A\mathbf{x}=0$.

**(a)** $A=\begin{bmatrix}4&2\\2&1\end{bmatrix}$.

**(b)** $A=\begin{bmatrix}2&1\\1&3\end{bmatrix}$.

## Proof.

(a) $A=\begin{bmatrix}4&2\\2&1\end{bmatrix}$.

Let $\mathbf{x}=\begin{bmatrix}x\\y\end{bmatrix}$ be a vector in $\mathbb{R}^2$. Then we have

$$\begin{aligned}\mathbf{x}^{\mathrm{T}}A\mathbf{x}&=\begin{bmatrix}x&y\end{bmatrix}\begin{bmatrix}4&2\\2&1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x&y\end{bmatrix}\begin{bmatrix}4x+2y\\2x+y\end{bmatrix}\\&=x(4x+2y)+y(2x+y)=4x^2+2xy+2xy+y^2\\&=4x^2+4xy+y^2\\&=(2x+y)^2\geq 0.\end{aligned}$$

Thus we have $\mathbf{x}^{\mathrm{T}}A\mathbf{x}\geq 0$.
Note that $\mathbf{x}^{\mathrm{T}}A\mathbf{x}=0$ if and only if $2x+y=0$.
Thus those vectors $\mathbf{x}$ such that $\mathbf{x}^{\mathrm{T}}A\mathbf{x}=0$ are

$$\mathbf{x}=\begin{bmatrix}x\\-2x\end{bmatrix}=x\begin{bmatrix}1\\-2\end{bmatrix}$$

for any real number $x$.

(b) $A=\begin{bmatrix}2&1\\1&3\end{bmatrix}$.

Let $\mathbf{x}=\begin{bmatrix}x\\y\end{bmatrix}$ be a vector in $\mathbb{R}^2$.
We compute

$$\begin{aligned}\mathbf{x}^{\mathrm{T}}A\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2x+y \\ x+3y \end{bmatrix} \\ &= x(2x+y)+y(x+3y) = 2x^2+xy+xy+3y^2 \\ &= 2x^2+2xy+3y^2 \\ &= x^2+(x+y)^2+2y^2 \geq 0. \end{aligned} \tag{*}$$

Thus we obtain $\mathbf{x}^{\mathrm{T}}A\mathbf{x} \geq 0$.
It follows from (*) that $\mathbf{x}^{\mathrm{T}}A\mathbf{x} = 0$ if and only if

$$x^2+(x+y)^2+2y^2=0.$$

Since each of $x^2, (x+y)^2, 2y^2$ is nonnegative, we must have $x=y=0$.
Therefore $\mathbf{x}^{\mathrm{T}}A\mathbf{x} = 0$ if and only if $\mathbf{x}=\mathbf{0}$.

---

35 **(a)** Prove that the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ does not have a square root. Namely, show that there is no complex matrix $B$ such that $B^2=A$.
**(b)** Prove that the $2 \times 2$ identity matrix $I$ has infinitely many distinct square root matrices.

Proof.

(a) The matrix $A$ does not have a square root.

Let $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be the matrix such that $B^2=A$.

Since

$$B^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{bmatrix} = \begin{bmatrix} a^2+bc & (a+d)b \\ (a+d)c & cb+d^2 \end{bmatrix},$$

it follows from $B^2=A$ that

$$\begin{bmatrix} a^2+bc & (a+d)b \\ (a+d)c & cb+d^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Comparing entries we obtain four equations

$$\begin{aligned} a^2+bc &= 0 && (1) \\ (a+d)b &= 1 && (2) \\ (a+d)c &= 0 && (3) \\ cb+d^2 &= 0. && (4) \end{aligned}$$

Equation (3) gives $a+d=0$ or $c=0$.
If $a+d=0$, then equation (2) becomes $0=1$. This is impossible and thus $c=0$.
Since $c=0$, equations (1) and (4) yield that $a=d=0$.
However, inserting these into (2) gives $0=1$.
Hence there is no solution satisfying these equations.

Therefore, there is no square root of the matrix $A$.

### (b) The identity matrix has infinitely many square roots

Let

$$B = \begin{bmatrix} 1 & r \\ 0 & -1 \end{bmatrix},$$

where $r$ be an arbitrary real numbers.

Then we compute directly and obtain

$$B^2 = \begin{bmatrix} 1 & r \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & r \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Hence $B$ is a square root of the identity matrix $I$.
Since $r$ is an arbitrary real numbers, there are infinitely many square roots of $I$.

---

36 Let $D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}$ be a diagonal matrix with distinct diagonal entries: $d_i \neq d_j$ if $i \neq j$. Let $A = (a_{ij})$ be an $n \times n$ matrix such that $A$ commutes with $D$, that is, $AD = DA$. Then prove that $A$ is a diagonal matrix.

### Proof.

We prove that the $(i, j)$-entry of $A$ is $a_{ij} = 0$ for $i \neq j$.

We compare the $(i, j)$-entries of both sides of $AD = DA$.
Let $D = (d_{ij})$. That is, $d_{ii} = d_i$ and $d_{ij} = 0$ if $i \neq j$.
The $(i, j)$-entry of $AD$ is

$$(AD)_{ij} = \sum_{k=1}^{n} a_{ik} d_{kj} = a_{ij} d_j.$$

The $(i, j)$-entry of $DA$ is

$$(DA)_{ij} = \sum_{k=1}^{n} d_{ik} a_{kj} = d_i a_{ij}.$$

Hence we have

$$a_{ij} d_j = a_{ij} d_i.$$

Or equivalently we have

$$a_{ij}(d_j - d_i) = 0.$$

Since $d_i \neq d_j$, we have $d_j - d_i \neq 0$.
Thus, we must have $a_{ij} = 0$ for $i \neq j$.

Hence $A = (a_{ij})$ is a diagonal matrix.

# Elementary Row Operations

**Definition**

Let $A$ be an $m \times n$ matrix.

① The following three operations on rows of a matrix are called **elementary row operations**.

1. **Interchanging two rows**:
   $R_i \leftrightarrow R_j$ interchanges rows $i$ and $j$.
2. **Multiplying a row by a non-zero scalar**:
   $tR_i$ multiplies row $i$ by the non-zero scalar (number) $t$.
3. **Adding a multiple of one row to another row**:
   $R_j + tR_i$ adds $t$ times row $i$ to row $j$.

② Two matrices are **row equivalent** if one can be obtained from the other by a sequence of elementary row operations.

③ The matrix in reduced row echelon form that is row equivalent to $A$ is denoted by $\mathrm{rref}(A)$.

④ The **rank** of a matrix $A$ is the number of rows in $\mathrm{rref}(A)$.

## Problems

① For each of the following matrices, find a row-equivalent matrix which is in reduced row echelon form. Then determine the rank of each matrix.

**(a)** $A = \begin{bmatrix} 1 & 3 \\ -2 & 2 \end{bmatrix}$.

**(b)** $B = \begin{bmatrix} 2 & 6 & -2 \\ 3 & -2 & 8 \end{bmatrix}$.

**(c)** $C = \begin{bmatrix} 2 & -2 & 4 \\ 4 & 1 & -2 \\ 6 & -1 & 2 \end{bmatrix}$.

**(d)** $D = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$.

**(e)** $E = \begin{bmatrix} -2 & 3 & 1 \end{bmatrix}$.

### Definition (Rank of a Matrix).

The **rank** of a matrix $A$ is the number of nonzero rows in the reduced row echelon form matrix rref($A$) that is row equivalent to $A$.

### Solution.

(a) $A = \begin{bmatrix} 1 & 3 \\ -2 & 2 \end{bmatrix}$

The matrix $A$ has rank 2, which can be seen by computing

$$\begin{bmatrix} 1 & 3 \\ -2 & 2 \end{bmatrix} \xrightarrow{R_2+2R_1} \begin{bmatrix} 1 & 3 \\ 0 & 8 \end{bmatrix} \xrightarrow{\frac{1}{8}R_2} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1-3R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Because the row-reduced matrix has two non-zero rows, the rank of $A$ is 2.

(b) $B = \begin{bmatrix} 2 & 6 & -2 \\ 3 & -2 & 8 \end{bmatrix}$

The matrix $B$ has rank 2, which can be seen by computing

$$\begin{bmatrix} 2 & 6 & -2 \\ 3 & -2 & 8 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & 3 & -1 \\ 3 & -2 & 8 \end{bmatrix} \xrightarrow{R_2-3R_1} \begin{bmatrix} 1 & 3 & -1 \\ 0 & -11 & 11 \end{bmatrix}$$

$$\xrightarrow{\frac{-1}{11}R_2} \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -1 \end{bmatrix} \xrightarrow{R_1-3R_2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix}$$

(c) $C = \begin{bmatrix} 2 & -2 & 4 \\ 4 & 1 & -2 \\ 6 & -1 & 2 \end{bmatrix}$

The matrix $C$ has rank 2, which can be seen by computing

$$\begin{bmatrix} 2 & -2 & 4 \\ 4 & 1 & -2 \\ 6 & -1 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & -2 \\ 6 & -1 & 2 \end{bmatrix} \xrightarrow[R_3-6R_1]{R_2-4R_1} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & -10 \\ 0 & 5 & -10 \end{bmatrix}$$

$$\xrightarrow{R_3-R_2} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & -10 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{5}R_2} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1+R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

(d) $D = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$

The matrix $D$ has rank 1, which can be seen by calculating:

$$\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} \xrightarrow{\frac{-1}{2}R_1} \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} \xrightarrow{R_2-3R_1,\, R_3-R_1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

(e) $E = \begin{bmatrix} -2 & 3 & 1 \end{bmatrix}$

The matrix $E$ has rank 1, which can be seen by calculating:

$$\begin{bmatrix} -2 & 3 & 1 \end{bmatrix} \xrightarrow{\frac{-1}{2}R_1} \begin{bmatrix} 1 & \frac{-3}{2} & \frac{-1}{2} \end{bmatrix}.$$

② Let $A$ and $I$ be $2 \times 2$ matrices defined as follows.

$$A = \begin{bmatrix} 1 & b \\ c & d \end{bmatrix}, \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Prove that the matrix $A$ is row equivalent to the matrix $I$ if $d - cb \neq 0$.

## Elementary row operations

We review the elementary row operations before the proof.

The three elementary row operations on a matrix are defined as follows.

③ Interchanging two rows:
$R_i \leftrightarrow R_j$ interchanges rows $i$ and $j$.

④ Multiplying a row by a non-zero scalar (a number):
$tR_i$ multiplies row $i$ by the non-zero scalar (number) $t$.

⑤ Adding a multiple of one row to another row:
$R_j + tR_i$ adds $t$ times row $i$ to row $j$.

## Proof.

Suppose that $b - cd \neq 0$. Then we can obtain the matrix $I$ from the matrix $A$ by the following sequence of elementary row operations.

First, we apply $R_2 - cR_1$ and get

$$A = \begin{bmatrix} 1 & b \\ c & d \end{bmatrix} \xrightarrow{R_2 - cR_1} \begin{bmatrix} 1 & b \\ 0 & d - cb \end{bmatrix}.$$

The next step is $\frac{1}{d-cb} R_2$ and get

$$\begin{bmatrix} 1 & b \\ 0 & d - cb \end{bmatrix} \xrightarrow{\frac{1}{d-cb} R_2} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}.$$

Note that this is where we need the assumption $d - cb \neq 0$ since $d - cb$ is in the denominator.

The last step is $R_1 - bR_2$, and we obtain

$$\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 - bR_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

In summary, we have the following sequence of elementary row operations from the matrix $A$ to the matrix I

$I.$

$$A = \begin{bmatrix} 1 & b \\ c & d \end{bmatrix} \xrightarrow{R_2 - cR_1} \begin{bmatrix} 1 & b \\ 0 & d-cb \end{bmatrix}$$

$$\xrightarrow{\frac{1}{d-cb} R_2} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 - bR_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I,$$

and hence $A$ and $I$ are row equivalent.

---

6 If $A, B$ have the same rank, can we conclude that they are row-equivalent? If so, then prove it. If not, then provide a counterexample.

## Solution.

Having the same rank does not mean they are row-equivalent.

For a simple counterexample, consider $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \end{bmatrix}$.

Both of these matrices have rank 1, but are not row-equivalent because they are already in reduced row echelon form.

## Another solution.

The problem doesn't specify the sizes of matrices $A, B$.

Note that if the sizes of $A$ and $B$ are distinct, then they can never be row-equivalent.
Keeping this in mind, let us consider the following two matrices.

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Then both matrices are in reduced row echelon form and have rank $1$.
As noted above, they are not row-equivalent because the sizes are distinct.

---

7 Find the rank of the following real matrix.

$$\begin{bmatrix} a & 1 & 2 \\ 1 & 1 & 1 \\ -1 & 1 & 1-a \end{bmatrix},$$

where $a$ is a real number.
(*Kyoto University*)

## Solution.

The rank is the number of nonzero rows of a (reduced) row echelon form matrix of the given matrix.

We apply elementary row operations as follows.

$$\begin{bmatrix} a & 1 & 2 \\ 1 & 1 & 1 \\ -1 & 1 & 1-a \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 1 \\ a & 1 & 2 \\ -1 & 1 & 1-a \end{bmatrix} \xrightarrow[R_3+R_1]{R_2-aR_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1-a & 2-a \\ 0 & 2 & 2-a \end{bmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2-a \\ 0 & 1-a & 2-a \end{bmatrix} \xrightarrow{R_3 - \frac{1-a}{2} R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2-a \\ 0 & 0 & (2-a)(a+1)/2 \end{bmatrix}.$$

The last matrix is in row echelon form.

Therefore, if $a \neq -1, 2$, then $(3,3)$-entry of the last matrix is not zero. From this we see that the rank is $3$ when $a \neq -1, 2$.

On the other hand, when $a = -1$ or $a = 2$ the third row is a zero row, hence the rank is $2$.

---

⑧ For an $m \times n$ matrix $A$, we denote by $\mathrm{rref}(A)$ the matrix in reduced row echelon form that is row equivalent to $A$. For example, consider the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix}$

Then we have

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2} R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

and the last matrix is in reduced row echelon form.

Hence $\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Find an example of matrices $A$ and $B$ such that

$$\mathrm{rref}(AB) \neq \mathrm{rref}(A)\mathrm{rref}(B).$$

### Proof.

Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then $A$ is already in reduced row echelon from.

So we have $\mathrm{rref}(A) = A$.

Applying the elementary row operations, we obtain

$$B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

As the last matrix is in reduced row echelon from, we have $\mathrm{rref}(B) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Therefore, we see that

$$\mathrm{rref}(A)\mathrm{rref}(B) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The product of $A$ and $B$ is

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

It follows that $\mathrm{rref}(AB) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

---

In summary, we have

$$\mathrm{rref}(AB) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathrm{rref}(A)\mathrm{rref}(B),$$

as required.

9 **(a)** Find all $3 \times 3$ matrices which are in reduced row echelon form and have rank 1.
**(b)** Find all such matrices with rank 2.

## Solution.

### (a) Find all $3 \times 3$ matrices which are in reduced row echelon form and have rank 1.

First we look at the rank 1 case.

For a $3 \times 3$ matrix in reduced row echelon form to have rank 1, it must have 2 rows which are all 0s. The non-zero row must be the first row, and it must have a leading 1.

These conditions imply that the matrix must be of one of the following forms:

$$\begin{bmatrix} 1 & a & b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & c \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

### (b) Find all such matrices with rank 2.

For a rank 2 $3 \times 3$ matrix in reduced row echelon form, there must be one row, the bottom one, which has only 0s.

Thus we need two leading 1s in distinct columns, and every other term in the same column with a leading 1 must be 0. The possibilities are:

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & a & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

10 If $A, B, C$ are three $m \times n$ matrices such that $A$ is row-equivalent to $B$ and $B$ is row-equivalent to $C$, then can we conclude that $A$ is row-equivalent to $C$? If so, then prove it. If not, then provide a counterexample.

## Definition (Row Equivalent).

Two matrices are said to be **row equivalent** if one can be obtained from the other by a sequence of elementary row operations.

Proof.

Yes, in this case $A$ and $C$ are row-equivalent.

By assumption, the matrices $A$ and $B$ are row-equivalent, which means that there is a sequence of elementary row operations that turns $A$ into $B$.

Call this sequence $r_1, r_2, \cdots, r_n$, where each $r_i$ is an elementary row operation.
(Start with applying $r_1$ to $A$.)

By another assumption, $B$ is row-equivalent to $C$, which means that there is a sequence of elementary row operations which transforms $B$ into $C$; call this sequence $s_1, s_2, \cdots, s_m$.

Putting these sequences together, the operations $r_1, r_2, \cdots, r_n$ , $s_1, s_2, \cdots, s_m$ will transform the matrix $A$ into $C$.

This proves that $A$ and $C$ are row-equivalent.

---

11 Prove that if $A$ is an $n \times n$ matrix with rank $n$, then $\operatorname{rref}(A)$ is the identity matrix.

Proof.

Because $A$ has rank $n$, we know that the $n \times n$ matrix rref($A$) has $n$ non-zero rows.
This means that all $n$ rows must have a leading 1.

Each leading 1 must be in a distinct column, so we must have that each of the $n$ columns has a leading 1.
In a row echelon matrix, these leading 1s must be arranged to lie on the diagonal.

In a reduced row echelon matrix, each column with a leading 1 has 0s above and below that 1.

These restrictions mean that rref($A$) must be the identity matrix.

---

12 Recall that a matrix $A$ is **symmetric** if $A^{\mathrm{T}} = A$, where $A^{\mathrm{T}}$ is the transpose of $A$. Is it true that if $A$ is a symmetric matrix and in reduced row echelon form, then $A$ is diagonal? If so, prove it. Otherwise, provide a counterexample.

Proof.

This is true.

If $A$ is in reduced row echelon form, then every term below the diagonal must be 0.
That is, the entry $a_{ij} = 0$ for all $i > j$.

If $A$ is, additionally, symmetric, then for $i < j$ we also have $a_{ji} = a_{ij} = 0$.

These two facts together means that $a_{ij} = 0$ whenever $i \neq j$.

In this case, the only non-zero terms in the matrix $A$ lie on the diagonal, and so $A$ is a diagonal matrix.

# Gaussian-Jordan Elimination

**Definition**

Consider the $m \times n$ system of linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

① The **coefficient matrix** of the system is

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

② The **augmented matrix** of the system is

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

③ [Gauss-Jordan Elimination] For a given system of linear equations, we can find a solution as follows. This procedure is called **Gauss-Jordan elimination**.

1. Write the augmented matrix of the system of linear equations.
2. Use elementaray row operations to reduce the augmented matrix into (reduced) row echelon form.
3. Write the system of linear equation corresponding to the matrix in row echelon form.
4. Solve the system using back substitution.

## Problems

① Solve the following system by transforming the augmented matrix to reduced echelon form. Indicate the elementary row operations you performed.

$$\begin{aligned} x_1 + x_2 - x_5 &= 1 \\ x_2 + 2x_3 + x_4 + 3x_5 &= 1 \\ x_1 - x_3 + x_4 + x_5 &= 0 \end{aligned}$$

**Solution.**

(a) Solve the following system by Gauss-Jordan elimination

The augmented matrix of the system is

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 1 & 0 & -1 & 1 & 1 & 0 \end{array}\right].$$

We apply elementary row operations as follows.

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 1 & 0 & -1 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_3-R_1} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 0 & -1 & -1 & 1 & 2 & -1 \end{array}\right] \xrightarrow{\substack{R_1-R_2 \\ R_3+R_2}}$$

$$\left[\begin{array}{ccccc|c} 1 & 0 & -2 & -1 & -4 & 0 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 5 & 0 \end{array}\right] \xrightarrow{\substack{R_1+2R_3 \\ R_2-2R_3}} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 3 & 6 & 0 \\ 0 & 1 & 0 & -3 & -7 & 1 \\ 0 & 0 & 1 & 2 & 5 & 0 \end{array}\right].$$

The last matrix is in reduced row echelon form.

The unknowns $x_1, x_2, x_3$ correspond to the leading $1$'s. Thus they are dependent variables and $x_4$ and $x_5$ are independent (free) variables.

From the last matrix, we see that the solution set is given by

$$\begin{aligned} x_1 &= -3x_4 - 6x_5 \\ x_2 &= 3x_4 + 7x_5 + 1 \\ x_3 &= -2x_4 - 5x_5 \end{aligned}$$

for any numbers $x_4, x_5$.

### (b) Determine all possibilities for the solution set of a homogeneous system

A homogeneous system is always consistent because it has the zero solution $x_1 = 0, x_2 = 0$. So, we already know the possibilities for the solution set are either a unique solution or infinitely many solutions.

Note that the homogeneous system we are inspecting has a solution $x_1 = 1, x_2 = 5$, which is different from the zero solution $x_1 = 0, x_2 = 0$. Thus, the homogeneous system has at least two solutions.

Thus, the only possibility is that the homogeneous system has infinitely many solutions.

② Solve the following system of linear equations using Gaussian elimination.

$$\begin{aligned} x + 2y + 3z &= 4 \\ 5x + 6y + 7z &= 8 \\ 9x + 10y + 11z &= 12 \end{aligned}$$

### Elementary row operations

The three *elementary row operations on a matrix* are defined as follows.

(1) **Interchanging two rows**:

$R_i \leftrightarrow R_j$ interchanges rows $i$ and $j$.

(2) **Multiplying a row by a non-zero scalar (a number):**

$tR_i$ multiplies row $i$ by the non-zero scalar (number) $t$.

(3) **Adding a multiple of one row to another row**:

$R_j + tR_i$ adds $t$ times row $i$ to row $j$.

## Solution.

First, the augmented matrix of the system is

$$A = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{array}\right].$$

We apply elementary row operations as follows to reduce the system to row echelon form.

$$A \xrightarrow{R_3-9R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & -8 & -16 & -24 \end{array}\right] \xrightarrow{-\frac{1}{8}R_3} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 1 & 2 & 3 \end{array}\right]$$

$$\xrightarrow{R_2-5R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 1 & 2 & 3 \end{array}\right] \xrightarrow{-\frac{1}{4}R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{array}\right]$$

$$\xrightarrow{R_3-R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The last matrix is in row echelon form.

The corresponding system of linear equations of it is

$$\begin{aligned} x + 2y + 3z &= 4 \\ y + 2z &= 3 \\ 0z &= 0 \end{aligned}$$

The last equation $0z = 0$ means that $z$ can be any number.
(More systematically, the variables corresponding to leading $1$'s in the echelon form matrix are dependent variables, and the rests are independent (free) variables.)

So let us say that $t$ is a value for $z$, namely $z = t$.
Then from the second equation, we have $y = -2t + 3$.
From the first equation, we have

$$x = -2y - 3z + 4 = -2(-2t + 3) - 3t + 4 = t - 2.$$

Thus the solution set is

$$(x, y, z) = (t - 2, -2t + 3, t)$$

for any $t$.

### Comment.

You may want to check whether the answer is correct by substituting this solution to the original equations.

Also, if you further reduce the matrix into *reduced* row echelon form, the last system becomes simpler (and simplest in a sense).
This procedure, to reduce a matrix until reduced row echelon form, is called the *Gauss-Jordan elimination*.

---

③ Solve the following system of linear equations using Gauss-Jordan elimination.

$$\begin{aligned} 6x + 8y + 6z + 3w &= -3 \\ 6x - 8y + 6z - 3w &= 3 \\ 8y \phantom{+6z} - 6w &= 6 \end{aligned}$$

We use the following notation.

**Elementary row operations.**

The three elementary row operations on a matrix are defined as follows.

(1) Interchanging two rows:

$R_i \leftrightarrow R_j$ interchanges rows $i$ and $j$.

(2) Multiplying a row by a non-zero scalar (a number):

$tR_i$ multiplies row $i$ by the non-zero scalar (number) $t$.

(3) Adding a multiple of one row to another row:

$R_j + tR_i$ adds $t$ times row $i$ to row $j$.

### Solution.

The augmented matrix of the system is

$$A = \left[\begin{array}{cccc|c} 6 & 8 & 6 & 3 & -3 \\ 6 & -8 & 6 & -3 & 3 \\ 0 & 8 & 0 & -6 & 6 \end{array}\right].$$

We apply elementary row operations as follows to reduce the system to a matrix in reduced row echelon form.

$$A \xrightarrow{R_2-R_1} \left[\begin{array}{cccc|c} 6 & 8 & 6 & 3 & -3 \\ 0 & -16 & 0 & -6 & 6 \\ 0 & 8 & 0 & -6 & 6 \end{array}\right] \xrightarrow[R_2+2R_3]{R_1-R_3} \left[\begin{array}{cccc|c} 6 & 0 & 6 & 9 & -9 \\ 0 & 0 & 0 & -18 & 18 \\ 0 & 8 & 0 & -6 & 6 \end{array}\right]$$

$$\xrightarrow[\frac{1}{2}R_3]{\frac{1}{3}R_1,\ \frac{-1}{18}R_2} \left[\begin{array}{cccc|c} 2 & 0 & 2 & 3 & -3 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 4 & 0 & -3 & 3 \end{array}\right] \xrightarrow[R_3+3R_2]{R_1-3R_2} \left[\begin{array}{cccc|c} 2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 4 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[\frac{1}{4}R_3]{\frac{1}{2}R_1} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

The last matrix is in reduced row echelon form.

The corresponding system of linear equations is

$$\begin{aligned} x+z&=0 \\ y&=0 \\ w&=-1 \end{aligned}$$

Let $z=t$ be a free variable. Then the solution is $(x,y,z,w)=(-t,0,t,-1)$ for any number $t$.

---

④ Solve the following system of linear equations by transforming its augmented matrix to reduced echelon form (Gauss-Jordan elimination). Find the vector form for the general solution.

$$\begin{aligned} x_1-x_3-3x_5&=1 \\ 3x_1+x_2-x_3+x_4-9x_5&=3 \\ x_1-x_3+x_4-2x_5&=1. \end{aligned}$$

**Solution.**

The augmented matrix of the given system is

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -3 & 1 \\ 3 & 1 & -1 & 1 & -9 & 3 \\ 1 & 0 & -1 & 1 & -2 & 1 \end{array}\right].$$

We apply the elementary row operations as follows.

We have

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -3 & 1 \\ 3 & 1 & -1 & 1 & -9 & 3 \\ 1 & 0 & -1 & 1 & -2 & 1 \end{array}\right] \xrightarrow{\substack{R_2-3R_1 \\ R_3-R_1}} \left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{R_2-R_3} \left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -3 & 1 \\ 0 & 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right].$$

The last matrix is in reduced row echelon form.

From this reduction, we see that the general solution is

$$
\begin{aligned}
x_1 &= x_3 + 3x_5 + 1 \\
x_2 &= -2x_3 + x_5 \\
x_4 &= -x_5.
\end{aligned}
$$

Here $x_3, x_5$ are free (independent) variables and $x_1, x_2, x_4$ are dependent variables.

To find the vector form for the general solution, we substitute these equations into the vector $\mathbf{x}$ as follows.

We have

$$
\begin{aligned}
\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} &= \begin{bmatrix} x_3 + 3x_5 + 1 \\ -2x_3 + x_5 \\ x_3 \\ -x_5 \\ x_5 \end{bmatrix} \\
&= \begin{bmatrix} x_3 \\ -2x_3 \\ x_3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 3x_5 \\ x_5 \\ 0 \\ -x_5 \\ x_5 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\
&= x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 3 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.
\end{aligned}
$$

Therefore **the vector form for the general solution** is given by

$$
\mathbf{x} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 3 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},
$$

where $x_3, x_5$ are free variables.

---

⑤ The given matrix is the augmented matrix for a system of linear equations. Give the vector form for the general solution.

$$
\left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -2 & 0 \\ 0 & 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right].
$$

Solution.

(a) Give the vector form for the general solution

The system of linear equation represented by the augmented matrix is

$$\begin{aligned} x_1 - x_3 - 2x_5 &= 0 \\ x_2 + 2x_3 - x_5 &= 0 \\ x_4 + x_5 &= 0. \end{aligned}$$

Solving the preceding system, we obtain

$$\begin{aligned} x_1 &= x_3 + 2x_5 \\ x_2 &= -2x_3 + x_5 \\ x_4 &= -x_5. \end{aligned}$$

Here $x_2, x_5$ are free variables and the rest are dependent variables.

Hence the general solution $\mathbf{x}$ can be expressed as

$$\begin{aligned} \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} &= \begin{bmatrix} x_3 + 2x_5 \\ -2x_3 + x_5 \\ x_3 \\ -x_5 \\ x_5 \end{bmatrix} \\ &= \begin{bmatrix} x_3 \\ -2x_3 \\ x_3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2x_5 \\ x_5 \\ 0 \\ -x_5 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}. \end{aligned}$$

Therefore, the vector form for the general solution is

$$\mathbf{x} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

### (b) Matrix product and transpose

First of all, note that by the property of the transpose we have

$$(A - B)^{\mathrm{T}} = A^{\mathrm{T}} - B^{\mathrm{T}}.$$

Hence we can simply the middle part:

$$\begin{aligned} A^{\mathrm{T}} - (A - B)^{\mathrm{T}} &= A^{\mathrm{T}} - (A^{\mathrm{T}} - B^{\mathrm{T}}) \\ &= A^{\mathrm{T}} - A^{\mathrm{T}} + B^{\mathrm{T}} = B^{\mathrm{T}}. \end{aligned}$$

Thus the expression becomes

$$\mathbf{v}^{\mathrm{T}} \left( A^{\mathrm{T}} - (A - B)^{\mathrm{T}} \right) C = \mathbf{v}^{\mathrm{T}} B^{\mathrm{T}} C.$$

The transposes of $\mathbf{v}$ and $B$ are

$$\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \text{ and } B^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Thus we have

$$\begin{aligned} \mathbf{v}^{\mathrm{T}} \left(A^{\mathrm{T}} - (A - B)^{\mathrm{T}}\right) C &= \mathbf{v}^{\mathrm{T}} B^{\mathrm{T}} C \\ &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 6 \end{bmatrix}. \end{aligned}$$

In conclusion, we have obtained

$$\mathbf{v}^{\mathrm{T}} \left(A^{\mathrm{T}} - (A - B)^{\mathrm{T}}\right) C = \begin{bmatrix} 0 & 6 \end{bmatrix}.$$

---

6 Solve the following system of linear equations and give the vector form for the general solution.

$$\begin{aligned} x_1 - x_3 - 2x_5 &= 1 \\ x_2 + 3x_3 - x_5 &= 2 \\ 2x_1 - 2x_3 + x_4 - 3x_5 &= 0 \end{aligned}$$

(*The Ohio State University*)

## Solution.

We solve the system by Gauss-Jordan elimination.

The augmented matrix of the system is given by

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 0 & -1 & 2 \\ 2 & 0 & -2 & 1 & -3 & 0 \end{array}\right].$$

We apply the elementary row operations as follows.

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 0 & -1 & 2 \\ 2 & 0 & -2 & 1 & -3 & 0 \end{array}\right] \xrightarrow{R_3 - 2R_1} \left[\begin{array}{ccccc|c} 1 & 0 & -1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 & 1 & -2 \end{array}\right].$$

Then the last matrix is in reduced row echelon form.

The variables $x_1, x_2, x_4$ correspond to the leading $1$'s of the last matrix, hence they are dependent variables and the rest $x_3, x_5$ are free variables.

From the last matrix we obtain the general solution

$$\begin{aligned} x_1 &= x_3 + 2x_5 + 1 \\ x_2 &= -3x_3 + x_5 + 2 \\ x_4 &= -x_5 - 2. \end{aligned}$$

The vector form for the general solution is obtained by substituting these into the vector $\mathbf{x}$. We have

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} x_3 + 2x_5 + 1 \\ -3x_3 + x_5 + 2 \\ x_3 \\ -x_5 - 2 \\ x_5 \end{bmatrix}$$

$$= x_3 \begin{bmatrix} 1 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 0 \\ -2 \\ 0 \end{bmatrix}.$$

Therefore, the vector form for the general solution is given by

$$\mathbf{x} = x_3 \begin{bmatrix} 1 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 0 \\ -2 \\ 0 \end{bmatrix},$$

where $x_3, x_5$ are free variables.

---

7 Determine whether the following augmented matrices are in reduced row echelon form, and calculate the solution sets of their associated systems of linear equations.

**(a)** $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 6 \end{array}\right].$

**(b)** $\left[\begin{array}{ccc|c} 1 & 0 & 3 & -4 \\ 0 & 1 & 2 & 0 \end{array}\right].$

**(c)** $\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 1 & -1 \end{array}\right].$

Solution.

(a) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 6 \end{array}\right]$

The first matrix is in reduced row echelon form.

For this matrix, we can read off the solution $x_1 = 2, x_2 = -3, x_3 = 6$.

(b) $\left[\begin{array}{ccc|c} 1 & 0 & 3 & -4 \\ 0 & 1 & 2 & 0 \end{array}\right]$

The second matrix is in reduced row echelon form.

For this matrix, the variable $x_3$ is a free variable because there are no leading 1s in the 3rd column.

The solution set can be expressed as $x_1 = -4 - 3x_3$, $x_2 = -2x_3$, and $x_3$ can be any real number.

(c) $\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 1 & -1 \end{array}\right]$

The third matrix is not in reduced echelon form because the bottom-left entry is $1$, not $0$, so we first use elementary row operations to put it in this form.

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 1 & -1 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & -1 & -1 \end{array}\right]$$
$$\xrightarrow{(-1)R_2} \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 1 & 1 \end{array}\right] \xrightarrow{R_1 - 2R_2} \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 1 \end{array}\right].$$

With the reduced matrix, we can read off the solution $x_1 = -2$ and $x_2 = 1$.

---

8 Find a cubic polynomial $p(x) = a + bx + cx^2 + dx^3$ such that $p(1) = 1, p'(1) = 5, p(-1) = 3$, and $p'(-1) = 1$.

**Solution.**

By differentiating $p(x)$, we obtain

$$p'(x) = b + 2cx + 3dx^2.$$

Thus the given conditions are

$$\begin{aligned} p(1) &= a + b + c + d = 1 \\ p'(1) &= b + 2c + 3d = 5 \\ p(-1) &= a - b + c - d = 3 \\ p'(-1) &= b - 2c + 3d = 1. \end{aligned}$$

We want to solve these four equations for $a, b, c, d$. So $a, b, c, d$ are unknowns and we regard the conditions as a system of four linear equations.
We solve the system using Gauss-Jordan elimination.

---

The augmented matrix for the system is

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 5 \\ 1 & -1 & 1 & -1 & 3 \\ 0 & 1 & -2 & 3 & 1 \end{array}\right]$$

We apply the elementary row operations as follows.

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 5 \\ 1 & -1 & 1 & -1 & 3 \\ 0 & 1 & -2 & 3 & 1 \end{array}\right] \xrightarrow{R_3-R_1} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 5 \\ 0 & -2 & 0 & -2 & 2 \\ 0 & 1 & -2 & 3 & 1 \end{array}\right] \xrightarrow{\substack{R_1-R_2 \\ R_3+2R_2 \\ R_4-R_2}}$$

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & -2 & -4 \\ 0 & 1 & 2 & 3 & 5 \\ 0 & 0 & 4 & 4 & 12 \\ 0 & 0 & -4 & 0 & -4 \end{array}\right] \xrightarrow[\frac{-1}{4}R_4]{\frac{1}{4}R_3} \left[\begin{array}{cccc|c} 1 & 0 & -1 & -2 & -4 \\ 0 & 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{\substack{R_1+R_4 \\ R_2-2R_4 \\ R_3-R_4}}$$

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & -3 \\ 0 & 1 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{R_3 \leftrightarrow R_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & -3 \\ 0 & 1 & 0 & 3 & 3 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right]$$

$$\xrightarrow{\substack{R_1+2R_4 \\ R_2-3R_4}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right].$$

The last matrix is in reduced row echelon form. Thus the solution to the system is

$$a = 1, b = -3, c = 1, d = 2.$$

Therefore, the polynomial is

$$p(x) = 1 - 3x + x^2 + 2x^3.$$

---

9 Find a quadratic function $f(x) = ax^2 + bx + c$ such that $f(1) = 3$, $f'(1) = 3$, and $f''(1) = 2$. Here, $f'(x)$ and $f''(x)$ denote the first and second derivatives, respectively.

**Solution.**

Each condition required on $f$ can be turned into an equation involving the constants $a, b, c$.

In particular, $f(1) = 3$ tells us that $a + b + c = 3$.

Because $f'(x) = 2ax + b$, the condition $f'(1) = 3$ gives us $2a + b = 3$.

And finally $f''(x) = 2a$, and so $f''(1) = 2a = 2$. Thus we have the system of equations

$$\begin{aligned} a + b + c &= 3 \\ 2a + b &= 3 \\ 2a &= 2 \end{aligned}$$

To solve this system, we could create the augmented matrix and then reduce it.

For this system, though, it is simpler to solve directly. The third equation tells us that $a = 1$.

Plugging this value into the second equation, we find $b = 1$.

Plugging both of these values into the first equation, we see $c = 1$ as well.

Thus the function we want is $f(x) = x^2 + x + 1$.

---

**10** **(a)** Find a function $g(\theta) = a\cos(\theta) + b\cos(2\theta) + c\cos(3\theta)$ such that $g(0) = g(\pi/2) = g(\pi) = 0$, where $a, b, c$ are constants.
**(b)** Find real numbers $a, b, c$ such that the function $g(\theta) = a\cos(\theta) + b\cos(2\theta) + c\cos(3\theta)$ satisfies $g(0) = 3$, $g(\pi/2) = 1$, and $g(\pi) = -5$.

## Solution.

### (a) Condition: $g(0) = g(\pi/2) = g(\pi) = 0$

Each condition required on $g$ can be turned into an equation involving the constants $a, b, c$. In particular, we have the system of linear equations

$$\begin{aligned} g(0) &= a + b + c = 0 \\ g\left(\frac{\pi}{2}\right) &= -b = 0 \\ g(\pi) &= -a + b - c = 0. \end{aligned}$$

---

To solve this system, we will use Gauss-Jordan elimination to reduce its augmented matrix.

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \end{array}\right] \xrightarrow[R_3+R_2]{R_1+R_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \end{array}\right] \xrightarrow[R_3+R_1]{(-1)R_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The solution can now be read off: $a + c = 0$ and $b = 0$.
Thus a general solution is of the form $g(\theta) = a\cos(\theta) - a\cos(3\theta)$, where $a$ is any real number.

### (b) Condition: $g(0) = 3$, $g(\pi/2) = 1$, and $g(\pi) = -5$

Each condition required on $g$ can be turned into an equation involving the constants $a, b, c$. In particular, we have

$$\begin{aligned} g(0) &= a + b + c = 3 \\ g(\pi/2) &= -b = 1 \\ g(\pi) &= -a + b - c = -5. \end{aligned}$$

---

To solve this system, we will use Gauss-Jordan elimination to reduce its augmented matrix.

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 3 \\ 0 & -1 & 0 & 1 \\ -1 & 1 & -1 & -5 \end{array}\right] \xrightarrow[R_3+R_2]{R_1+R_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 4 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & -1 & -4 \end{array}\right] \xrightarrow[R_3+R_1]{(-1)R_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The solution can now be read off: $a + c = 4$ and $b = -1$. Thus a general solution is of the form

$$g(\theta) = a\cos(\theta) - \cos(2\theta) + (4 - a)\cos(3\theta),$$

where $a$ is any real number.

---

⑪ A 2-digit number has two properties: The digits sum to 11, and if the number is written with digits reversed, and subtracted from the original number, the result is 45. Find the number.

**Solution.**

The key to this problem is noticing that our 2-digit number can be written as $10A + B$, where $A, B$ are the 10s and 1s digit respectively.

The digits summing to 11 then yields the equality $A + B = 11$.

The number with digits reversed is $10B + A$, and so the second property yields the equation $10A + B – (10B + A) = 45$. Simplifying, we have the system of equations

$$\begin{aligned} A + B &= 11 \\ 9A - 9B &= 45 \end{aligned}$$

To solve this system, we create the augmented matrix and use elementary row operations to put it into reduced row-echelon form:

$$\left[\begin{array}{cc|c} 1 & 1 & 11 \\ 9 & -9 & 45 \end{array}\right] \xrightarrow{R_2 - 9R_1} \left[\begin{array}{cc|c} 1 & 1 & 11 \\ 0 & -18 & -54 \end{array}\right]$$

$$\xrightarrow{\frac{-1}{18}R_2} \left[\begin{array}{cc|c} 1 & 1 & 11 \\ 0 & 1 & 3 \end{array}\right] \xrightarrow{R_1 - R_2} \left[\begin{array}{cc|c} 1 & 0 & 8 \\ 0 & 1 & 3 \end{array}\right].$$

With this final augmented matrix, we can read off the solution $A = 8$ and $B = 3$.

Thus our number is $10A + B = 83$.

# Consistent and Inconsistent Systems of Linear Equations

**Definition**

1. An $m \times n$ **system of linear equations** is

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= b_3 \\ \vdots \qquad\qquad \cdots \qquad\qquad &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned} \tag{*}$$

where $x_1, x_2, \dots, x_n$ are unknowns (variables) and $a_{ij}, b_k$ are numbers.
Thus an $m \times n$ system of linear equations consists of $m$ equations and $n$ unknowns $x_1, x_2, \dots, x_n$.

2. A system of linear equations is called **homogeneous** if the constants $b_1, b_2, \dots, b_m$ are all zero.

3. A **solution** of the system (*) is a sequence of numbers $s_1, s_2, \dots, s_n$ such that the substitution $x_1 = s_1, x_2 = s_2, \dots, x_n = s_n$ satisfies all the $m$ equations in the system (*).

4. The rank of a system of linear equation is the rank of the coefficient matrix.

**Summary**

1. For a given system of linear equations, there are only three possibilities for the solution set of the system: No solution (inconsistent), a unique solution, or infinitely many solutions.

2. The possibilities for the solution set of a homogeneous system is either a unique solution or infinitely many solutions.

3. If $m < n$, then an $m \times n$ system is either inconsistent or it has infinitely many solutions.

4. If $m < n$, then an $m \times n$ homogeneous system has infinitely many solutions.

5. If a consistent $m \times n$ system of linear equation has rank $r$, then the system has $n - r$ free variables.

## Problems

1. Determine all possibilities for the solution set of the system of linear equations described below.
   - **(a)** A homogeneous system of $3$ equations in $5$ unknowns.
   - **(b)** A homogeneous system of $5$ equations in $4$ unknowns.
   - **(c)** A system of $5$ equations in $4$ unknowns.
   - **(d)** A system of $2$ equations in $3$ unknowns that has $x_1 = 1, x_2 = -5, x_3 = 0$ as a solution.
   - **(e)** A homogeneous system of $4$ equations in $4$ unknowns.
   - **(f)** A homogeneous system of $3$ equations in $4$ unknowns.
   - **(g)** A homogeneous system that has $x_1 = 3, x_2 = -2, x_3 = 1$ as a solution.
   - **(h)** A homogeneous system of $5$ equations in $3$ unknowns and the rank of the system is $3$.

**(i)** A system of $3$ equations in $2$ unknowns and the rank of the system is $2$.
**(j)** A homogeneous system of $4$ equations in $3$ unknowns and the rank of the system is $2$.

## the possibilities for the solution set of a system of linear equations

An $m \times n$ **system of linear equations** is

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= b_3 \\ \vdots \qquad\qquad \cdots \qquad\qquad &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

where $x_1, x_2, \ldots, x_n$ are unknowns (variables) and $a_{ij}, b_k$ are numbers.

Thus an $m \times n$ system of linear equations consists of $m$ equations and $n$ unknowns $x_1, x_2, \ldots, x_n$.

A system of linear equations is called **homogeneous** if the constants $b_1, b_2, \ldots, b_m$ are all zero.

Namely, a homogeneous system is

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= 0 \\ \vdots \qquad\qquad \cdots \qquad\qquad &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

A **solution** of the system (*) is a sequence of numbers $s_1, s_2, \ldots, s_n$ such that the substitution $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ satisfies all the $m$ equations in the system (*).

We sometimes use the vector notation and say

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

is a solution of the system.

For example, every homogeneous system has the **zero solution** $x_1 = 0, x_2 = 0, \ldots, x_n = 0$, or

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Here we summarize several theorems concerning with the possibilities for the number of solutions of a system of linear equations.

We say a system is **consistent** if the system has at least one solution.

A system is called **inconsistent** if the system has no solutions at all.

**Theorem 1**. For a given system of linear equations, there are three possibilities for the solution set of the system: No solution (inconsistent), a unique solution, or infinitely many solutions.

Thus, for example, if we find two distinct solutions for a system, then it follows from the theorem that there are infinitely many solutions for the system.

Next, since a homogeneous system has the zero solution, it is always consistent. Thus:

**Theorem 2**. The possibilities for the solution set of a homogeneous system is either a unique solution or infinitely many solutions.

Let us refine these theorems. Suppose that an $m \times n$ system of linear equations is given. That is, there are $m$ linear equations and $n$ unknowns.

**Theorem 3**. If $m < n$, then the system is either inconsistent or it has infinitely many solutions.

Thus, there are only two possibilities when $m < n$: No solution or infinitely many solutions.

**Theorem 4**. Consider $m \times n$ homogeneous system of linear equations. Then the system has always infinitely many solutions.

This is obtained by noting that a homogeneous system always has the zero solution, hence consistent. By the previous theorem, the only possibility is infinitely many solutions.

### Summary 1

**Summary 1**: The number of solutions of a system of linear equations is one of 0, 1, or $\infty$. If there are more unknowns ($n$) than the number of equations ($m$), then the number of solutions of the system is either 0 or $\infty$. If a system is homogeneous, then it has the zero solution and thus a homogeneous system is always consistent.

## The case $m \geq n$?

What happens when $m \geq n$?
In general, when the number of equations is greater than or equal to the number of unknowns, we cannot narrow down the possibilities.
We need more information about the system. The key word is the **rank** of the system.
For a given system (*), let $A$ be the coefficient matrix and let $\mathbf{b}$ be the constant term vector. Then we define the rank of the system to be the rank of the augmented matrix $[A \mid \mathbf{b}]$.
Recall that the rank is defined as follows. We first reduce the matrix $[A \mid \mathbf{b}]$ to a matrix $[A' \mid \mathbf{b}']$ in

(reduced) row echelon form by elementary row operations.

Then the rank of $[A \mid \mathbf{b}]$ is the number of nonzero rows in the matrix $[A' \mid \mathbf{b}']$.

**Theorem 5**. Consider an $m \times n$ system of linear equations. Suppose that it is consistent. Then the rank $r$ of the system satisfies $r \leq n$. Also, the system has $n - r$ free variables.

A free variable means an unknown that can be assigned arbitrary values. It follows that if a system has a free variable, then there are infinitely many solutions.

**Caution**: the theorem assumes that a given system is consistent.

### Summary 2

**Summary 2**: Suppose that an $m \times n$ system of linear equations is consistent and let $r$ be the rank of the system. If $n = r$, then the system has a unique solution. If $n > r$, then the system has infinitely many solutions.

## Problems and solutions

Determine all possibilities for the solution set of the system of linear equations described below.

**(a)** A homogeneous system of 3 equations in 5 unknowns.

**(b)** A homogeneous system of 5 equations in 4 unknowns.

**(c)** A system of 5 equations in 4 unknowns.

**(d)** A system of 2 equations in 3 unknowns that has $x_1 = 1, x_2 = -5, x_3 = 0$ as a solution.

**(e)** A homogeneous system of 4 equations in 4 unknowns.

**(f)** A homogeneous system of 3 equations in 4 unknowns.

**(g)** A homogeneous system that has $x_1 = 3, x_2 = -2, x_3 = 1$ as a solution.

**(h)** A homogeneous system of 5 equations in 3 unknowns and the rank of the system is 3.

**(i)** A system of 3 equations in 2 unknowns and the rank of the system is 2.

**(j)** A homogeneous system of 4 equations in 3 unknowns and the rank of the system is 2.

## Solutions.

In the solution, $m$ denotes the number of equations and $n$ denotes the number of unknowns in the given system.

### (a) A homogeneous system of 3 equations in 5 unknowns.

Since the system is homogeneous, it has the zero solution, hence consistent. Since there are more unknowns than equations, there are infinitely many solutions.

(b) A homogeneous system of 5 equations in 4 unknowns.

Since the system is homogeneous, it has the zero solution. Since there are more equations than unknowns, we cannot determine further.
Thus the possibilities are either a unique solution or infinitely many solution.
(If the rank $r$ of the system is 4, then a unique solution. If $r < 4$, then there are infinitely many solutions.)

(c) A system of 5 equations in 4 unknowns.

Since $m > n$, we can only say that the possibilities are no solution, a unique solution, or infinitely many solution.

(d) A system of 2 equations in 3 unknowns that has $x_1 = 1, x_2 = -5, x_3 = 0$ as a solution.

Since $m < n$, the system is either inconsistent or has infinitely many solutions. Since $x_1 = 1, x_2 = -5, x_3 = 0$ is a solution of the system, the system is not inconsistent. Thus the only possibility is infinitely many solutions.

(e) A homogeneous system of 4 equations in 4 unknowns.

Since $m = n$, this tells nothing. But since the system is homogeneous it has the zero solution, hence consistent. The possibilities are either a unique solution or infinitely many solutions.

(f) A homogeneous system of 3 equations in 4 unknowns.

Since $m < n$, the system has no solution or infinitely many solutions. But a homogeneous system is always consistent. Thus, the only possibility is infinitely many solutions.

(g) A homogeneous system that has $x_1 = 3, x_2 = -2, x_3 = 1$ as a solution.

The possibilities for the solution set for any homogeneous system is either a unique solution or infinitely many solutions. Since the homogeneous system has the zero solution and $x_1 = 3, x_2 = -2, x_3 = 1$ is another solution, it has at least two distinct solution. Thus the only possibility is infinitely many solutions.

(h) A homogeneous system of 5 equations in 3 unknowns and the rank of the system is 3.

A homogeneous system is always consistent. Since the rank $r$ of the system and the number $n$ of unknowns are equal, the only possibility is the zero solution (and the zero solution is a unique solution).

(i) A system of 3 equations in 2 unknowns and the rank of the system is 2.

We don't know whether the system is consistent or not.
If it is consistent, then since the rank $r$ and the number of unknowns are the same, the system has a unique solution. Thus the possibilities are either inconsistent or a unique solution.

Before talking about the rank, we need to discuss whether the system is inconsistent or not. For example, consider the following $3 \times 2$ system

$$\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 3 \\ 3x_1 + 3x_2 &= 3. \end{aligned}$$

Then the augmented matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}.$$

We reduce this by elementary row operations as follows.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix} \xrightarrow{\substack{R_2-2R_1 \\ R_3-3R_1}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is in echelon form and it has two nonzero rows. Thus, the rank of the system is 2. However, the second row means that we have $0 = 1$. Hence the system is inconsistent.

(j) A homogeneous system of 4 equations in 3 unknowns and the rank of the system is 2.

A homogeneous system is consistent. The rank is $r = 2$ and the number of variables is $n = 3$. Hence there is $n - r = 1$ free variable. Thus there are infinitely many solutions.

2 Determine all possibilities for the number of solutions of each of the system of linear equations described below.
**(a)** A system of $5$ equations in $3$ unknowns and it has $x_1 = 0, x_2 = -3, x_3 = 1$ as a solution.
**(b)** A homogeneous system of $5$ equations in $4$ unknowns and the rank of the system is $4$.
(*The Ohio State University*)

## Hint.

See the post Summary: possibilities for the solution set of a system of linear equations to review how to determine the number of solutions of a system of linear equations.

### Solution.

Let $m$ be the number of equations and $n$ be the number of unknowns in the given system.
Note that the information $m > n$ tells us nothing about the possibilities for the number of solutions.

### (a) A system of $5$ equations in $3$ unknowns and it has $x_1 = 0, x_2 = -3, x_3 = 1$ as a solution.

The system has at least one solution $x_1 = 0, x_2 = -3, x_3 = 1$, hence it is consistent.
Thus, the system has either a unique solution or infinitely many solutions. Since $m > n$, we cannot narrow down the possibilities further. Thus, the possibilities are either a unique solution (which must be $x_1 = 0, x_2 = -3, x_3 = 1$) or infinitely many solution.

### (b) A homogeneous system of $5$ equations in $4$ unknowns and the rank of the system is $4$.

Since the system is homogeneous, it has always the zero solution, hence it is consistent.
The fact $m > n$ gives no new information. But since the rank $r$ of the system is equal to the number of unknowns $n$, there is no free variable. ($n - r = 0$.) Thus, the system must have a unique solution, which is the zero solution.

### Comment.

This is one of the midterm exam 1 problems of linear algebra (Math 2568) at the Ohio State University.

Some students wrongly concluded from the condition $m > n$. Again, note that the condition $m > n$ gives no new information at all.

③ Determine all possibilities for the solution set of a homogeneous system of $2$ equations in $2$ unknowns that has a solution $x_1 = 1, x_2 = 5$.

### (b) Determine all possibilities for the solution set of a homogeneous system

A homogeneous system is always consistent because it has the zero solution $x_1 = 0, x_2 = 0$.
So, we already know the possibilities for the solution set are either a unique solution or infinitely many solutions.

Note that the homogeneous system we are inspecting has a solution $x_1 = 1, x_2 = 5$, which is different from the zero solution $x_1 = 0, x_2 = 0$. Thus, the homogeneous system has at least two solutions.
Thus, the only possibility is that the homogeneous system has infinitely many solutions.

---

④ For what value(s) of $a$ does the system have nontrivial solutions?

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ -x_1 - x_2 + x_3 &= 0 \\ 3x_1 + 4x_2 + ax_3 &= 0. \end{aligned}$$

### Solution.

First note that the system is homogeneous and hence it is consistent. Thus if the system has a nontrivial solution, then it has infinitely many solutions.

This happens if and only if the system has at least one free variable. The number of free variables is $n-r$, where $n$ is the number of unknowns and $r$ is the rank of the augmented matrix.

To find the rank, we reduce the augmented matrix by elementary row operations.

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ -1 & -1 & 1 & 0 \\ 3 & 4 & a & 0 \end{array}\right] \xrightarrow[R_3-3R_1]{R_2+R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -2 & a-3 & 0 \end{array}\right]$$
$$\xrightarrow{R_3+2R_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & a+1 & 0 \end{array}\right].$$

The last matrix is in row echelon form.
Thus if $a+1=0$, then the third row is a zero row, hence the rank is $2$. In this case we have $n-r=3-2=1$ free variable. Thus there are infinitely many solutions. In particular, the system has nontrivial solutions.

On the other hand, if $a+1\neq 0$, then the rank is $3$ and there is no free variables since $n-r=3-3=0$.

In summary, the system has nontrivial solutions exactly when $a=-1$.

5 Suppose that the following matrix $A$ is the augmented matrix for a system of linear equations.

$$A=\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & -1 & -2 & a^2 \\ -1 & -7 & -11 & a \end{array}\right],$$

where $a$ is a real number. Determine all the values of $a$ so that the corresponding system is consistent.

## Solution.

We apply the elementary row operations to $A$ as follows.

$$A=\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & -1 & -2 & a^2 \\ -1 & -7 & -11 & a \end{array}\right] \xrightarrow{\substack{R_2-2R_1 \\ R_3+R_1}} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -5 & -8 & a^2-8 \\ 0 & -5 & -8 & a+4 \end{array}\right]$$
$$\xrightarrow{R_3-R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -5 & -8 & a^2-8 \\ 0 & 0 & 0 & -a^2+a+12 \end{array}\right].$$

The last matrix is in row echelon form. (It's not in reduced row echelon form.)
Then the system is consistent if and only if $-a^2+a+12$ is zero. (See Remark below.)
Since we can factor

$$0 = -a^2 + a + 12 = -(a+3)(a-4),$$

we see that the system is consistent if and only if $a = -3, 4$.

Remark

If $-a^2 + a + 12 \neq 0$, then we divide the third row by $-a^2 + a + 12$ and obtain

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 1 \end{array}\right]$$

in the third row.

The corresponding linear equation is

$$0x_1 + 0x_2 + 0x_3 = 1,$$

and clearly, there is no solution to this equation. Hence the system is inconsistent.

On the other hand, if $-a^2 + a + 12 = 0$, then the system has solutions. (You just need to reduce the above matrix further or use the back substitution.)

---

6 Determine whether the following systems of equations (or matrix equations) described below has no solution, one unique solution or infinitely many solutions and justify your answer.

**(a)**

$$\begin{cases} ax + by = c \\ dx + ey = f, \end{cases}$$

where $a, b, c, d$ are scalars satisfying $a/d = b/e = c/f$.

**(b)** $A\mathbf{x} = \mathbf{0}$, where $A$ is a non-singular matrix.

**(c)** A homogeneous system of $3$ equations in $4$ unknowns.

**(d)** $A\mathbf{x} = \mathbf{b}$, where the row-reduced echelon form of the augmented matrix $[A|\mathbf{b}]$ looks as follows:

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(*The Ohio State University*)

Hint.

Recall that possibilities for the solution set of a system of linear equation is either no solution (inconsistent) or one unique solution or infinitely many solutions.

A homogeneous system is a system with zero constant terms.

A homogeneous system always has the zero solution.

Solution.

**(a)** Note that by the given condition $a/d = b/e = c/f$, the numbers $b, e, c$ are not zero. The augmented matrix of the system is

$$\left[\begin{array}{cc|c} a & b & c \\ d & e & f \end{array}\right].$$

We apply the elementary row operations as follows.

$$\left[\begin{array}{cc|c} a & b & c \\ d & e & f \end{array}\right] \xrightarrow{R_1-\frac{a}{d}R_2} \left[\begin{array}{cc|c} 0 & 0 & 0 \\ d & e & f \end{array}\right] \xrightarrow{R_1\leftrightarrow R_2} \left[\begin{array}{cc|c} d & e & f \\ 0 & 0 & 0 \end{array}\right]$$
$$\xrightarrow{\frac{1}{d}} \left[\begin{array}{cc|c} 1 & e/d & f/d \\ 0 & 0 & 0 \end{array}\right].$$

Here in the first step, we used the relation $a/d = b/e = c/f$.
The last matrix is in reduced row echelon form with rank $1$ and it does not have a row of the form $[00|1]$. Therefore the system is consistent and there are $1(= \text{the number of unknowns} \ - \ \text{rank})$ free variable, hence there are infinitely many solutions.

**(b)** The system is homogeneous, thus it has the zero solution. Since the coefficient matrix $A$ is singular, the system has non-zero solution as well.
Therefore, the only possibility is that the system has infinitely many solutions.

**(c)** A homogeneous system has the zero solution hence it is consistent. Since there are more unknowns than equations, the system must have infinitely many solutions.

**(d)** The last row of the reduced row echelon form matrix of the augmented matrix is $[000|1]$.
It corresponds to the equation

$$0x_1 + 0x_2 + 0x_3 = 1.$$

Equivalently, this is $0 = 1$ and this is impossible. Thus the system has no solution (an inconsistent system).

# Linear Combination and Linear Independence

**Definition**

1. The expression $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ is called a **linear combination** of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, where $c_1, c_2, \ldots, c_k$ are scalars in $\mathbb{R}$.
2. A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is said to be **linearly independent** if the only scalrs $c_1, c_2, \ldots, c_k$ satisfying $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ are $c_1 = c_2 = \cdots = c_k = 0$. We also say that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent.
3. If vectors are not linearly independent, they are **linearly dependent**.

**Summary**

1. The set of $n$-dimensional vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ are linearly dependent if $k > n$. (If there are more vectors than the dimension, then the vectors are linearly dependent.)

## Problems

1. Express the vector $\mathbf{b} = \begin{bmatrix} 2 \\ 13 \\ 6 \end{bmatrix}$ as a linear combination of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}.$$

(*The Ohio State University, Linear Algebra Exam*)

### Solution.

We need to find numbers $x_1, x_2, x_3$ satisfying

$$x_1\mathrm{v}_1 + x_2\mathrm{v}_2 + x_3\mathrm{v}_3 = \mathrm{b}.$$

This vector equation is equivalent to the following matrix equation.

$$[\mathrm{v}_1\mathrm{v}_2\mathrm{v}_3]\mathrm{x} = \mathrm{b}$$

or more explicitly we can write it as

$$\begin{bmatrix} 1 & 1 & 1 \\ 5 & 2 & 4 \\ -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 13 \\ 6 \end{bmatrix}.$$

Thus the problem is to find the solution of this matrix equation.

Let us consider the augmented matrix for this system to apply Gauss-Jordan elimination.

The augmented matrix is

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 5 & 2 & 4 & 13 \\ -1 & 1 & 3 & 6 \end{array}\right].$$

We apply elementary row operations and obtain a matrix in reduced row echelon form as follows.

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 5 & 2 & 4 & 13 \\ -1 & 1 & 3 & 6 \end{array}\right] \xrightarrow[R_3+R_1]{R_2-5R_1} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 3 \\ 0 & 2 & 4 & 8 \end{array}\right]$$

$$\xrightarrow[-R_2]{\frac{1}{2}R_3} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & 3 & 1 & -3 \\ 0 & 1 & 2 & 4 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 1 & -3 \end{array}\right]$$

$$\xrightarrow[R_3-3R_2]{R_1-R_2} \left[\begin{array}{rrr|r} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & -5 & -15 \end{array}\right] \xrightarrow{\frac{-1}{5}R_3} \left[\begin{array}{rrr|r} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$$\xrightarrow[R_2-2R_3]{R_1+R_3} \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Therefore the solution for the system is

$$x_1 = 1, x_2 = -2, \text{ and } x_3 = 3$$

and we obtain the linear combination

$$\mathrm{b} = \mathrm{v}_1 - 2\mathrm{v}_2 + 3\mathrm{v}_3.$$

② Write the vector $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ as a linear combination of the vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}.$

**Solution.**

We want to find real numbers $x_1, x_2, x_3$ which solve the equation

$$\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 2x_3 \\ -2x_2 \\ x_2 + 4x_3 \end{bmatrix}.$$

Each row in this matrix now gives an equality, and so the three rows define a system of linear equations. To find a solution, we reduce its augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 2 & 1 \\ 0 & -2 & 0 & 3 \\ 0 & 1 & 4 & -1 \end{array}\right] \xrightarrow[R_2+2R_3]{R_1-2R_3} \left[\begin{array}{ccc|c} 1 & 0 & -6 & 3 \\ 0 & 0 & 8 & 1 \\ 0 & 1 & 4 & -1 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & 0 & -6 & 3 \\ 0 & 1 & 4 & -1 \\ 0 & 0 & 8 & 1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{8}R_3} \left[\begin{array}{ccc|c} 1 & 0 & -6 & 3 \\ 0 & 1 & 4 & -1 \\ 0 & 0 & 1 & \frac{1}{8} \end{array}\right] \xrightarrow[R_2-4R_3]{R_1+6R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{15}{4} \\ 0 & 1 & 0 & \frac{-3}{2} \\ 0 & 0 & 1 & \frac{1}{8} \end{array}\right].$$

Now we can read off the solution $x_1 = \frac{15}{4}$, $x_2 = \frac{-3}{2}$, and $x_3 = \frac{1}{8}$.

③ For what value(s) of $a$ is the following set $S$ linearly dependent?

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ a \end{bmatrix}, \begin{bmatrix} a \\ 0 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ a^2 \\ 7 \end{bmatrix}, \begin{bmatrix} 1 \\ a \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 3 \\ a^3 \end{bmatrix} \right\}.$$

## Solution.

### (a) Linearly dependent vectors

Since the set $S$ consists of five $4$-dimensional vectors, it is linearly dependent regardless of the value of $a$. Thus, for any value of $a$ the set $S$ is linearly dependent.

### (b) Perpendicular nonzero vectors are linearly independent

Suppose that we have the linear combination

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = 0 \tag{*}$$

for some scalars $c_1, c_2, c_3$.
To show that the set is linearly independent, we must show that $c_1 = c_2 = c_3 = 0$.

Taking the dot product with $\mathbf{v}_1$, we have

$$\mathbf{v}_1 \cdot (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) = \mathbf{v}_1 \cdot 0 = 0.$$

Distributing $\mathbf{v}_1$, we have

$$c_1\mathbf{v}_1 \cdot \mathbf{v}_1 + c_2\mathbf{v}_1 \cdot \mathbf{v}_2 + c_3\mathbf{v}_1 \cdot \mathbf{v}_3 = 0.$$

Since the dot products $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0, \mathbf{v}_1 \cdot \mathbf{v}_3 = 0$, we have

$$c_1||\mathbf{v}_1||^2 = c_1\mathbf{v}_1 \cdot \mathbf{v}_1 = 0.$$

Since the vector $\mathbf{v}_1 \neq 0$, the length $||\mathbf{v}|| \neq 0$.
Thus, we must have $c_1 = 0$.

We repeat the above argument with $\mathrm{v}_1$ replaced by $\mathrm{v}_2$ as follows.

Since $c_1 = 0$, the (*) becomes

$$c_2\mathrm{v}_2 + c_3\mathrm{v}_3 = 0.$$

We have

$$\begin{aligned} 0 &= \mathrm{v}_2 \cdot 0 = \mathrm{v}_2 \cdot (c_2\mathrm{v}_2 + c_3\mathrm{v}_3) \\ &= c_2\mathrm{v}_2 \cdot \mathrm{v}_2 + c_3\mathrm{v}_2 \cdot \mathrm{v}_3. \end{aligned}$$

Since $\mathrm{v}_2 \cdot \mathrm{v}_3 = 0$, we have

$$0 = c_2\mathrm{v}_2 \cdot \mathrm{v}_2 = c_2||\mathrm{v}_2||^2.$$

Since $\mathrm{x} \neq 0$, we have $||\mathrm{v}_2|| \neq 0$, and thus $c_2 = 0$.

Hence (*) is now just $c_3\mathrm{v}_3 = 0$, and since $\mathrm{x}_3 \neq 0$, we conclude that $c_3 = 0$. Therefore, we have shown that $c_1 = c_2 = c_3 = 0$, and thus the set $\{\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3\}$ is linearly independent.

4 Prove that any set of vectors which contains the zero vector is linearly dependent.

### Solution.

Let $\mathbf{0}$ be the zero vector, and $\mathbf{v}_1, \cdots, \mathbf{v}_k$ are the other vectors in the set.

Then we have the non-trivial linear combination

$$1 \cdot \mathbf{0} + 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k = \mathbf{0}.$$

This is a non-trivial linear combination because one of the coefficients is non-zero.

Thus by definition, the set $\{\mathbf{0}, \mathbf{v}_1, \cdots, \mathbf{v}_k\}$ is linearly dependent.

5 Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a set of nonzero vectors in $\mathbb{R}^m$ such that the dot product $mathbfv_i \cdot \mathbf{v}_j = 0$ when $i \neq j$. Prove that the set is linearly independent.

### Solution.

#### (a) Linearly dependent vectors

Since the set $S$ consists of five $4$-dimensional vectors, it is linearly dependent regardless of the value of $a$. Thus, for any value of $a$ the set $S$ is linearly dependent.

#### (b) Perpendicular nonzero vectors are linearly independent

Suppose that we have the linear combination

$$c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + c_3\mathrm{v}_3 = 0 \qquad (*)$$

for some scalars $c_1, c_2, c_3$.

To show that the set is linearly independent, we must show that $c_1 = c_2 = c_3 = 0$.

Taking the dot product with $\mathbf{v}_1$, we have

$$\mathbf{v}_1 \cdot (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) = \mathbf{v}_1 \cdot \mathbf{0} = 0.$$

Distributing $\mathbf{v}_1$, we have

$$c_1\mathbf{v}_1 \cdot \mathbf{v}_1 + c_2\mathbf{v}_1 \cdot \mathbf{v}_2 + c_3\mathbf{v}_1 \cdot \mathbf{v}_3 = 0.$$

Since the dot products $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0, \mathbf{v}_1 \cdot \mathbf{v}_3 = 0$, we have

$$c_1||\mathbf{v}_1||^2 = c_1\mathbf{v}_1 \cdot \mathbf{v}_1 = 0.$$

Since the vector $\mathbf{v}_1 \neq \mathbf{0}$, the length $||\mathbf{v}|| \neq 0$.
Thus, we must have $c_1 = 0$.

We repeat the above argument with $\mathbf{v}_1$ replaced by $\mathbf{v}_2$ as follows.
Since $c_1 = 0$, the (*) becomes

$$c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}.$$

We have

$$\begin{aligned} 0 = \mathbf{v}_2 \cdot \mathbf{0} &= \mathbf{v}_2 \cdot (c_2\mathbf{v}_2 + c_3\mathbf{v}_3) \\ &= c_2\mathbf{v}_2 \cdot \mathbf{v}_2 + c_3\mathbf{v}_2 \cdot \mathbf{v}_3. \end{aligned}$$

Since $\mathbf{v}_2 \cdot \mathbf{v}_3 = 0$, we have

$$0 = c_2\mathbf{v}_2 \cdot \mathbf{v}_2 = c_2||\mathbf{v}_2||^2.$$

Since $\mathbf{x} \neq \mathbf{0}$, we have $||\mathbf{v}_2|| \neq 0$, and thus $c_2 = 0$.

Hence (*) is now just $c_3\mathbf{v}_3 = \mathbf{0}$, and since $\mathbf{x}_3 \neq \mathbf{0}$, we conclude that $c_3 = 0$. Therefore, we have shown that $c_1 = c_2 = c_3 = 0$, and thus the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

⑥ Determine whether the following set of vectors is linearly independent or linearly dependent. If the set is linearly dependent, express one vector in the set as a linear combination of the others.

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \\ 7 \\ 11 \end{bmatrix} \right\}.$$

## Solution.

Consider the linear combination

$$x_1 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -2 \\ 7 \\ 11 \end{bmatrix} = \mathbf{0} \qquad (*)$$

with variables $x_1, x_2, x_3, x_4$.

We determine whether there is $(x_1, x_2, x_3, x_4) \neq (0,0,0,0)$ satisfying the linear combination (*).

The linear combination (*) is written as the matrix equation

$$\begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 2 & -2 & -2 \\ -1 & 3 & 0 & 7 \\ 0 & 4 & 1 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0.$$

---

To find the solutions of the equation, we apply the Gauss-Jordan elimination.

The augmented matrix $[A \mid 0]$ can be reduced by elementary row operations as follows.

$$[A \mid 0] = \left[\begin{array}{cccc|c} 1 & 1 & -1 & -2 & 0 \\ 0 & 2 & -2 & -2 & 0 \\ -1 & 3 & 0 & 7 & 0 \\ 0 & 4 & 1 & 11 & 0 \end{array}\right] \xrightarrow[\frac{1}{2}R_2]{R_3+R_1} \left[\begin{array}{cccc|c} 1 & 1 & -1 & -2 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 4 & -1 & 5 & 0 \\ 0 & 4 & 1 & 11 & 0 \end{array}\right]$$

$$\xrightarrow{\substack{R_1-R_2 \\ R_3-4R_2 \\ R_4-4R_2}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 3 & 9 & 0 \\ 0 & 0 & 5 & 15 & 0 \end{array}\right] \xrightarrow{\substack{\frac{1}{3}R_3 \\ \frac{1}{5}R_4}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right]$$

$$\xrightarrow{\substack{R_2+R_3 \\ R_4-R_3}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, the general solution is given by

$$\begin{aligned} x_1 &= x_4 \\ x_2 &= -2x_4 \\ x_3 &= -3x_4, \end{aligned}$$

where $x_4$ is a free variable.

---

If we let $x_4 = 1$, then we have a nonzero solution $x_1 = 1, x_2 = -2, x_3 = -3, x_4 = 1$.

Thus the set is linearly dependent.

Substituting these values into (*), we have

$$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - 3\begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ -2 \\ 7 \\ 11 \end{bmatrix} = 0.$$

Solving this for the last vector we obtain the linear combination

$$\begin{bmatrix}-2\\-2\\7\\11\end{bmatrix}=-\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}+2\begin{bmatrix}1\\2\\3\\4\end{bmatrix}+3\begin{bmatrix}-1\\-2\\0\\1\end{bmatrix}.$$

7 Find the value(s) of $h$ for which the following set of vectors

$$\left\{\mathbf{v}_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\mathbf{v}_2=\begin{bmatrix}h\\1\\-h\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}1\\2h\\3h+1\end{bmatrix}\right\}$$

is linearly independent.
(*Boston College*)

## Solution 1.

Let us consider the linear combination

$$x_1\mathrm{v}_1+x_2\mathrm{v}_2+x_3\mathrm{v}_3=0. \tag{*}$$

If this homogeneous system has only zero solution $x_1=x_2=x_3=0$, then the vectors $\mathrm{v}_1,\mathrm{v}_2,\mathrm{v}_3$ are linearly independent.
We reduce the augmented matrix for the system as follows.

$$\left[\begin{array}{ccc|c}1&h&1&0\\0&1&2h&0\\0&-h&3h+1&0\end{array}\right]\xrightarrow{R_3+hR_2}\left[\begin{array}{ccc|c}1&h&1&0\\0&1&2h&0\\0&0&2h^2+3h+1&0\end{array}\right].$$

From this, we see that the homogeneous system (*) has only the zero solution if and only if

$$2h^2+3h+1\neq 0.$$

Since we have

$$2h^2+3h+1=(2h+1)(h+1),$$

if $h\neq-\frac{1}{2},-1$, then $2h^2+3h+1\neq 0$.

In summary, the vectors $\mathrm{v}_1,\mathrm{v}_2,\mathrm{v}_3$ are linearly independent for any $h$ except $-\frac{1}{2},-1$.

## Solution 2.

Note that the vectors $\mathrm{v}_1,\mathrm{v}_2,\mathrm{v}_3$ are linearly independent if and only the matrix

$$A:=[\mathrm{v}_1\mathrm{v}_2\mathrm{v}_3]=\begin{bmatrix}1&h&1\\0&1&2h\\0&-h&3h+1\end{bmatrix}$$

is nonsingular.

Also, the matrix $A$ is nonsingular if and only if the determinant $\det(A)$ is nonzero.
So we compute the determinant of the matrix $A$ by the first column cofactor expansion and obtain

$$\begin{aligned}\det(A) &= \begin{vmatrix} 1 & h & 1 \\ 0 & 1 & 2h \\ 0 & -h & 3h+1 \end{vmatrix} \\ &= \begin{vmatrix} 1 & 2h \\ -h & 3h+1 \end{vmatrix} \\ &= 2h^2 + 3h + 1 \\ &= (2h+1)(h+1).\end{aligned}$$

Hence $\det(A) \neq 0$ if and only if $h \neq -\frac{1}{2}, -1$.

Therefore, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent for any values of $h$ except $-\frac{1}{2}, -1$.

8 Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ a \\ 5 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 4 \\ b \end{bmatrix}$$

be vectors in $\mathbb{R}^3$. Determine a condition on the scalars $a, b$ so that the set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.

## Solution.

Consider the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{0},$$

where $\mathbf{0}$ is the three-dimensional zero vector.

Our goal is to find a condition on $a$, $b$ so that the above equation has a nontrivial solution $x_1, x_2, x_3$.

This equation is equivalent to the 3 × 3 homogeneous system of linear equations

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & a & 4 \\ 0 & 5 & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}.$$

---

We solve the system by Gauss-Jordan elimination.

The augmented matrix of the system is

$$\left[\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 2 & a & 4 & 0 \\ 0 & 5 & b & 0 \end{array}\right].$$

Applying elementary row operations, we have

$$\begin{bmatrix}1&1&0&0\\2&a&4&0\\0&5&b&0\end{bmatrix}\xrightarrow{R_2-2R_1}\begin{bmatrix}1&1&0&0\\0&a-2&4&0\\0&5&b&0\end{bmatrix}\xrightarrow{\frac{1}{5}R_3}$$

$$\begin{bmatrix}1&1&0&0\\0&a-2&4&0\\0&1&b/5&0\end{bmatrix}\xrightarrow{R_2\leftrightarrow R_3}\begin{bmatrix}1&1&0&0\\0&1&b/5&0\\0&a-2&4&0\end{bmatrix}$$

$$\xrightarrow{R_3-(a-2)R_2}\begin{bmatrix}1&1&0&0\\0&1&b/5&0\\0&0&4-\frac{b(a-2)}{5}&0\end{bmatrix}$$

If the (3, 3) entry $4-\frac{b(a-2)}{5}$ of the last matrix is $4-\frac{b(a-2)}{5}$ is zero, then we obtain a matrix in echelon form

$$\begin{bmatrix}1&1&0&0\\0&1&b/5&0\\0&0&0&0\end{bmatrix}.$$

This implies that $x_3$ is a free variable, hence the homogeneous system has a nonzero solution $x_1, x_2, x_3$. Hence in this case the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.

On the other hand, if $4-\frac{b(a-2)}{5}\neq 0$, we divide the third row by this number and obtain

$$\begin{bmatrix}1&1&0&0\\0&1&b/5&0\\0&0&1&0\end{bmatrix},$$

and from this we see that the solution is $x_1 = x_2 = x_3 = 0$.
Thus in this case the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

In conclusion, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent if and only if $4-\frac{b(a-2)}{5}=0$.
Thus, the condition on $a, b$ is

$$b(a-2)=20.$$

### Related Question.

**Problem.**

Determine conditions on the scalars $a, b$ so that the following set $S$ of vectors is linearly dependent.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\},$$

where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ a \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 2 \\ b \end{bmatrix}.$$

⑨ Determine conditions on the scalars $a, b$ so that the following set $S$ of vectors is linearly dependent.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\},$$

where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ a \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 2 \\ b \end{bmatrix}.$$

## Solution.

Let us consider the linear combination

$$x_1\mathrm{v}_1 + x_2\mathrm{v}_2 + x_3\mathrm{v}_3 = 0.$$

We want to find conditions on $a, b$ so that there is $(x_1, x_2, x_3) \neq (0, 0, 0)$ satisfying this linear combination.

Since the linear combination can be written as

$$A\mathrm{x} = 0,$$

where

$$A = [\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3] = \begin{bmatrix} 1 & 1 & 0 \\ 3 & a & 2 \\ 1 & 4 & b \end{bmatrix}, \text{ and } \mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

our goal is the same as finding conditions on $a, b$ so that $A$ is a singular matrix.

---

We apply elementary row operations to the augmented matrix $[A \mid 0]$.

$$[A \mid 0]=\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 3 & a & 2 & 0 \\ 1 & 4 & b & 0 \end{array}\right]$$

$$\xrightarrow[R_3-R_1]{R_2-3R_1}\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & a-3 & 2 & 0 \\ 0 & 3 & b & 0 \end{array}\right]\xrightarrow{R_2\leftrightarrow R_3}\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 3 & b & 0 \\ 0 & a-3 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{\frac{1}{3}R_2}\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & b/3 & 0 \\ 0 & a-3 & 2 & 0 \end{array}\right]\xrightarrow{R_3-(a-3)R_2}\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & b/3 & 0 \\ 0 & 0 & 2-(a-3)b/3 & 0 \end{array}\right].$$

The last matrix is in echelon form.

---

Let $c=2-(a-3)b/3$ be the number in $(3,3)$-entry.
If $c\neq 0$, then it is easy to see that we can further reduce the matrix to

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

and the only solution is $x_1=x_2=x_3=0$.

---

On the other hand, if $c=0$, then the last row becomes a zero row and thus we see that $x_3$ is a free variable. Hence the system has a nonzero solution.

Therefore, the vectors $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ are linearly dependent if and only if $c=0$.
Hence the condition we are looking for is

$$c=2-(a-3)b/3=0,$$

or equivalently,

$$(a-3)b=6.$$

## Related Question.

**Problem.**
Let

$$\mathbf{v}_1=\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix},\mathbf{v}_2=\begin{bmatrix} 1 \\ a \\ 5 \end{bmatrix},\mathbf{v}_3=\begin{bmatrix} 0 \\ 4 \\ b \end{bmatrix}$$

be vectors in $\mathbb{R}^3$.

Determine a condition on the scalars $a, b$ so that the set of vectors $\{\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3\}$ is linearly dependent.

The solution is given in the post ↴

Determine a Condition on $a, b$ so that Vectors are Linearly Dependent

**10** **(a)** Prove that the column vectors of every $3 \times 5$ matrix $A$ are linearly dependent.
**(b)** Prove that the row vectors of every $5 \times 3$ matrix $B$ are linearly dependent.

## Proof.

### (a) Prove that the column vectors of every $3 \times 5$ matrix $A$ are linearly dependent.

Note that the column vectors of the matrix $A$ are linearly dependent if the matrix equation

$$A\mathrm{x} = 0$$

has a nonzero solution $\mathbf{x} \in \mathbb{R}^5$.

The equation is equivalent to a $3 \times 5$ homogeneous system.
As there are more variables than equations, the homogeneous system has infinitely many solutions.

In particular, the equation has a nonzero solution $\mathbf{x}$.
Hence the column vectors are linearly dependent.

### (b) Prove that the row vectors of every $5 \times 3$ matrix $B$ are linearly dependent.

Observe that the row vectors of the matrix $B$ are the column vectors of the transpose $B^{\mathrm{T}}$. Note that the size of $B^{\mathrm{T}}$ is $3 \times 5$.

In part (a), we showed that the column vectors of any $3 \times 5$ matrix are linearly dependent.
It follows that the column vectors of $B^{\mathrm{T}}$ are linearly dependent.
Hence the row vectors of $B$ are linearly dependent.

**11** Suppose $M$ is an $n \times n$ upper-triangular matrix. If the diagonal entries of $M$ are all non-zero, then prove that the column vectors are linearly independent. Does the conclusion hold if we do not assume that $M$ has non-zero diagonal entries?

## Proof.

### If the diagonal entries of $M$ are all non-zero, then prove that the column vectors are linearly independent.

Let $\mathbf{x}$ denote an arbitrary column vector of length $n$, and let $0$ denote the zero vector of the same size.

The columns of $M$ are linearly independent if and only if the only solution to the equation $M\mathbf{x} = 0$ is the vector $\mathbf{x} = 0$.

Consider the upper-triangular matrix

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & m_{1,3} & \cdots & m_{1,n} \\ 0 & m_{2,2} & m_{2,3} & \cdots & m_{2,n} \\ 0 & 0 & m_{3,3} & \cdots & m_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & m_{n,n} \end{bmatrix}.$$

The equation $M\mathbf{x} = 0$ then yields a system of linear equations with $n$ equations and $n$ variables. To find a solution, consider the augmented matrix $\left[ [c|c]M \quad 0 \right]$.

Because $M$ is upper-triangular, we can use back-substitution to solve. The bottom row of the augmented matrix gives the equation $m_{n,n} x_n = 0$.
By assumption, $m_{n,n} \neq 0$ because it is a diagonal entry. Thus we must have that $x_n = 0$.

Next, the second-to-last row in the augmented matrix gives the equation $m_{n-1,n-1} x_{n-1} + m_{n-1,n} x_n = 0$. Because $x_n = 0$ and $m_{n-1,n-1} \neq 0$, we must have that $x_{n-1} = 0$.

We continue working backward in this way to see that $x_i = 0$ for all $1 \leq i \leq n$. Thus $\mathbf{x} = 0$, and so the columns of $M$ must be linearly independent.

> Does the conclusion hold if we do not assume that $M$ has non-zero diagonal entries?

If the diagonal entries of $M$ could be non-zero, then the columns might be linearly dependent. Consider the simple example

$$M = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

⑫ Suppose that an $n \times m$ matrix $M$ is composed of the column vectors $\mathbf{b}_1, \cdots, \mathbf{b}_m$. Prove that a vector $\mathbf{v} \in \mathbb{R}^n$ can be written as a linear combination of the column vectors if and only if there is a vector $\mathbf{x}$ which solves the equation $M\mathbf{x} = \mathbf{v}$.

### Proof.

Suppose that $\mathrm{v}$ is a linear combination of the vectors $\mathrm{b}_1, \ldots, \mathrm{v}_m$. That is, suppose there are coefficients $x_1, x_2, \cdots, x_m$ such that

$$x_1 \mathrm{b}_1 + x_2 \mathrm{b}_2 + \cdots + x_m \mathrm{b}_m = \mathrm{v}.$$

Then define the column vector

$$\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}.$$

Then $\mathrm{x}$ satisfies the equation $M\mathrm{x} = \mathrm{v}$ .

---

On the other hand, suppose that there is a vector $\mathrm{x}$ which satisfies the desired equation with components $x_1, x_2, \cdots, x_m$ .

Then these components can be used to find the linear combination

$$x_1 \mathrm{b}_1 + x_2 \mathrm{b}_2 + \cdots + x_m \mathrm{b}_m = M\mathrm{x} = \mathrm{v}.$$

Hence, the vector $\mathrm{v}$ is a linear combination of column vectors $\mathrm{b}_1, \ldots, \mathrm{b}_m$ .

# Nonsingular Matrices

**Definition**

Let $A$ be an $n \times n$ matrix.

1. $A$ is nonsingular if the only solution to $A\mathbf{x} = \mathbf{0}$ is the zero solution $\mathbf{x} = \mathbf{0}$.

**Summary**

Let $A$ be an $n \times n$ matrix.

1. If $A$ is nonsingular, then $A^{\mathrm{T}}$ is nonsingular.
2. $A$ is nonsingular if and only if the column vectors of $A$ are linearly independent.
3. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $n \times 1$ column vector $\mathbf{b}$ if and only if $A$ is nonsingular.

## Problems

1. Determine whether the following matrices are nonsingular or not.

**(a)** $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix}$.

**(b)** $B = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 4 & 1 & 4 \end{bmatrix}$.

### Solution.

Recall that an $n \times n$ matrix is nonsingular if $Ax = 0$ has only the zero solution. This is equivalent to the condition that the rank of $A$ is $n$.

(a) Is $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix}$ nonsingular?

We find the rank of the matrix $A$ as follows.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix} \xrightarrow[R_3 - R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \xrightarrow{-\frac{1}{2}R_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 - R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The last matrix is in reduced row echelon form and has three nonzero rows.

Hence, the rank of $A$ is $3$.

It follows that the matrix $A$ is nonsingular.

(b) Is $B = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 4 & 1 & 4 \end{bmatrix}$ nonsingular?

Next, we compute the rank of $B$ as follows.

$$B = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 4 & 1 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 4 & 1 & 4 \end{bmatrix} \xrightarrow[R_3-4R_1]{R_2-2R_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{R_3-R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This computation shows that the rank of $B$ is $2$.

Hence, the matrix $B$ is singular.

② Consider the matrix $M = \begin{bmatrix} 1 & 4 \\ 3 & 12 \end{bmatrix}$.

**(a)** Show that $M$ is singular.

**(b)** Find a non-zero vector $\mathbf{v}$ such that $M\mathbf{v} = \mathbf{0}$, where $\mathbf{0}$ is the $2$-dimensional zero vector.

Solution.

(a) Show that $M$ is singular.

To see that $M$ is singular, we will row-reduce the augmented matrix:

$$\left[\begin{array}{cc|c} 1 & 4 & 0 \\ 3 & 12 & 0 \end{array}\right] \xrightarrow{R_2-3R_1} \left[\begin{array}{cc|c} 1 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The row-reduced matrix has a row of all $0$'s, and hence is singular.

(b) Find a non-zero vector $\mathbf{v}$ such that $M\mathbf{v} = \mathbf{0}$.

We see that the vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ must satisfy the equation $v_1 + 4v_2 = 0$.

There are many solutions. One solution, for example, is found by setting $v_1 = 1$ and solve $v_2 = -\frac{1}{4}$.

---

③ Let $A$ be the following $3 \times 3$ matrix.

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & 1 & a \end{bmatrix}.$$

Determine the values of $a$ so that the matrix $A$ is nonsingular.

Solution.

We use the fact that a matrix is nonsingular if and only if it is row equivalent to the identity matrix.

We apply the elementary row operations as follows.

$$A \xrightarrow{R_3-R_1} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & a+1 \end{bmatrix} \xrightarrow{R_1-R_2} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & a+1 \end{bmatrix}.$$

If $a+1=0$, then the last matrix is in reduced row echelon form.
Thus $A$ is not row equivalent to the identity matrix.

On the other hand, if $a+1 \neq 0$, then we can continue the reduction as follows.

$$\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & a+1 \end{bmatrix} \xrightarrow{\frac{1}{a+1}R_3} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_2-2R_3]{R_1+3R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore $A$ is row equivalent to the identity matrix.

We conclude that the matrix $A$ is nonsingular for any values of $a$ except for $a=-1$.

## Comment.

If you know how to compute the determinant of a $3 \times 3$ matrix, then you may also solve this using the fact that a matrix is nonsingular if and only if the determinant of it is nonzero.

---

4

Determine the values of a real number $a$ such that the matrix $A = \begin{bmatrix} 3 & 0 & a \\ 2 & 3 & 0 \\ 0 & 18a & a+1 \end{bmatrix}$ is nonsingular.

---

## Solution.

We apply elementary row operations and obtain:

$$A = \begin{bmatrix} 3 & 0 & a \\ 2 & 3 & 0 \\ 0 & 18a & a+1 \end{bmatrix} \xrightarrow{R_1-R_2} \begin{bmatrix} 1 & -3 & a \\ 2 & 3 & 0 \\ 0 & 18a & a+1 \end{bmatrix}$$

$$\xrightarrow{R_2-2R_1} \begin{bmatrix} 1 & -3 & a \\ 0 & 9 & -2a \\ 0 & 18a & a+1 \end{bmatrix} \xrightarrow{R_3-(2a)R_2} \begin{bmatrix} 1 & -3 & a \\ 0 & 9 & -2a \\ 0 & 0 & 4a^2+a+1 \end{bmatrix}.$$

From this, we see that the matrix $A$ is nonsingular if and only if the $(3,3)$-entry $4a^2+a+1$ is not zero.
By the quadratic formula, we see that

$$a = \frac{-1 \pm \sqrt{-15}}{8}$$

are solutions of $4a^2+a+1=0$.

Note that these are not real numbers. Thus, for any real number $a$, we have $4a^2+a+1 \neq 0$.

Hence, we can divide the third row by this number, and eventually we can reduce it to the identity matrix. So the rank of $A$ is $3$, and $A$ is nonsingular for any real number $a$.

⑤ **(a)** Suppose that a $3 \times 3$ system of linear equations is inconsistent. Is the coefficient matrix of the system nonsingular?
**(b)** Suppose that a $3 \times 3$ homogeneous system of linear equations has a solution $x_1 = 0, x_2 = -3, x_3 = 5$. Is the coefficient matrix of the system nonsingular?
**(c)** Let $A$ be a $4 \times 4$ matrix and let $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$. Suppose that we have $A\mathbf{v} = A\mathbf{w}$. Is the matrix $A$ nonsingular?

### Hint.

Recall the following fact.

Let $A$ be an $n \times n$ matrix.

Then $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $n \times 1$ column vector $\mathbf{b}$ if and only if $A$ is nonsingular.

### Solution.

> (a) Suppose that a $3 \times 3$ system of linear equations is inconsistent. Is the coefficient matrix of the system nonsingular?

Let $A\mathbf{x} = \mathbf{b}$ be the system, where $A$ is the coefficient matrix and $\mathbf{b}$ is the constant term vector. Because this system is inconsistent, then $A$ is singular. (If $A$ would be nonsingular, then the system has a unique solution.)

> (b) Suppose that a $3 \times 3$ homogeneous system of linear equations has a solution $x_1 = 0, x_2 = -3, x_3 = 5$. Is the coefficient matrix of the system nonsingular?

Recall that a homogeneous system of linear equations has always the zero solution. As the system has another solution $x_1 = 0, x_2 = -3, x_3 = 5$, the system $A\mathbf{x} = \mathbf{0}$ must have infinitely many solutions. Here $A$ is the coefficient matrix.
Thus, the coefficient matrix $A$ is singular. (If $A$ would be nonsingular, the system has only one solution, which must be the zero solution.)

> (c) Suppose that we have $A\mathbf{v} = A\mathbf{w}$. Is the matrix $A$ nonsingular?

Because $A\mathbf{v} = A\mathbf{w}$, we have

$$A(\mathbf{v} - \mathbf{w}) = A\mathbf{v} - A\mathbf{w} = \mathbf{0}.$$

Note that

$$\mathbf{v} - \mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$

is a nonzero vector.

This implies that the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has infinitely many solutions as we have found a nonzero solution.

Hence, $A$ is singular.

---

6 Let $A$ be a $3 \times 3$ singular matrix. Then show that there exists a **nonzero** $3 \times 3$ matrix $B$ such that

$$AB = O,$$

where $O$ is the $3 \times 3$ zero matrix.

---

## Proof.

Since $A$ is singular, the equation $A\mathbf{x} = \mathbf{0}$ has a nonzero solution.

Let $\mathbf{x}_1$ be a nonzero solution of $A\mathbf{x} = \mathbf{0}$.

We define the $3 \times 3$ matrix $B$ to be

$$B = [\mathbf{x}_1, \mathbf{0}, \mathbf{0}],$$

that is, the first column of $B$ is the vector $x_1$ and the second and the third column vectors are $3$-dimensional zero vectors.

Then since $x_1 \neq \mathbf{0}$, the matrix $B$ is not the zero matrix.

We have

$$\begin{aligned} AB &= A[\mathbf{x}_1, \mathbf{0}, \mathbf{0}] \\ &= [A\mathbf{x}_1, A\mathbf{0}, A\mathbf{0}] \\ &= [\mathbf{0}, \mathbf{0}, \mathbf{0}] = O, \end{aligned}$$

since $A\mathbf{x}_1 = \mathbf{0}$.

Thus we have found the nonzero matrix $B$ such that the product $AB = O$.

## Comment.

This is one of the midterm exam 1 problems of linear algebra (Math 2568) at the Ohio State University.

In the exam the following hint was given:

Hint: Let $B = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$, where $\mathbf{x}_i$ is the $i$-th column vector of $B$ for $i = 1, 2, 3$. Then $AB = [A\mathbf{x}_1, A\mathbf{x}_2, A\mathbf{x}_3]$ .

## Common Mistake

Several students wrote “Since $A$ is singular, $A$ has a solution”.

This does not make sense. A solution of what? You need to remember the definition correctly.

A matrix $A$ is **singular** if the equation $A\mathbf{x} = \mathbf{0}$ has a nonzero solution $\mathbf{x}$.

Also, note that $\mathbf{x}$ is a vector, not a number.

⑦ Let $A$ be an $n \times n$ singular matrix. Then prove that there exists a nonzero $n \times n$ matrix $B$ such that $AB = O$, where $O$ is the $n \times n$ zero matrix.

**Definition.**

Recall that an $n \times n$ matrix $A$ is called **singular** if the equation

$$A\mathbf{x} = \mathbf{0}$$

has a nonzero solution $\mathbf{x} \in \mathbb{R}^n$.

**Proof.**

Since $A$ is singular, there exists a nonzero vector $\mathbf{b} \in \mathbb{R}^n$ such that

$$A\mathbf{b} = \mathbf{0}.$$

Then we define the $n \times n$ matrix $B$ whose first column is the vector $\mathbf{b}$ and the other entries are zero. That is

$$B = \begin{bmatrix} \mathbf{b} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}.$$

Since the vector $\mathbf{b}$ is nonzero, the matrix $B$ is nonzero.

With this choice of $B$, we have

$$\begin{aligned} AB &= A\begin{bmatrix} \mathbf{b} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{b} & A\mathbf{0} & \cdots & A\mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} = O. \end{aligned}$$

Hence we have proved that $AB = O$ with the nonzero matrix $B$.

---

⑧ Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be 2-dimensional vectors and let $A$ be a $2 \times 2$ matrix.
**(a)** Show that if $\mathbf{v}_1, \mathbf{v}_2$ are linearly dependent vectors, then the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ are also linearly dependent.
**(b)** If $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent vectors, can we conclude that the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ are also linearly independent?
**(c)** If $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent vectors and $A$ is nonsingular, then show that the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ are also linearly independent.

**Proof.**

(a) If $\mathbf{v}_1, \mathbf{v}_2$ are linearly dependent, then $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly dependent.

Since $\mathbf{v}_1, \mathbf{v}_2$ are linearly dependent vectors, there exists $(c_1, c_2) \neq (0, 0)$ such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}.$$

Using this equality, we obtain

$$\begin{aligned}\mathbf{0} &= A\mathbf{0} \\ &= A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) \\ &= A(c_1\mathbf{v}_1) + A(c_2\mathbf{v}_2) \\ &= c_1(A\mathbf{v}_1) + c_2(A\mathbf{v}_2) \qquad \text{(since } c_1, c_2 \text{ are scalars).}\end{aligned}$$

Hence we have

$$c_1(A\mathbf{v}_1) + c_2(A\mathbf{v}_2) = \mathbf{0}$$

with $(c_1, c_2) \neq (0, 0)$.
This implies that the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly dependent.

> (b) If $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent, can we conclude that $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly independent?

The answer is no. In general, even though $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent vectors, the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ might be linearly dependent. Let us give an example.

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
Then it is straightforward to see that these vectors are linearly independent.
Let

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

be the $2 \times 2$ zero matrix.

Then we have

$$A\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and clearly $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly dependent vectors.

> (c) If $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent and $A$ is nonsingular, then $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly independent.

Consider the linear combination

$$c_1(A\mathbf{v}_1) + c_2(A\mathbf{v}_2) = \mathbf{0}.$$

To show that $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly independent, we show that $c_1 = c_2 = 0$.

The above equality can be written as

$$A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = \mathbf{0}.$$

Since the matrix $A$ is nonsingular, the homogeneous equation $A\mathbf{x} = \mathbf{0}$ has only the zero solution.
Therefore we have

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}.$$

(Another reasoning here is that since the matrix $A$ is nonsingular, it is invertible, and thus we have the inverse matrix $A^{-1}$. Multiplying by $A^{-1}$ on the left, we obtain the same equation.)

Now, since $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent by assumption, it follows that $c_1 = c_2 = 0$.

Hence we conclude that the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly independent.

⑨ Let $A$ be an $n \times n$ matrix. Suppose that the sum of elements in each row of $A$ is zero. Then prove that the matrix $A$ is singular.

## Definition.

An $n \times n$ matrix $A$ is said to be **singular** if there exists a nonzero vector $\mathrm{v}$ such that $A\mathrm{v} = 0$.

Otherwise, we say that $A$ is a **nonsingular** matrix.

## Proof.

Let $\mathrm{v} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ be the $n$-dimensional vector all of whose entries are $1$.

Then we compute $A\mathrm{v}$.

Note that $A\mathrm{v}$ is an $n$-dimensional vector, and the $i$-th entry of $A\mathrm{v}$ is the sum of the $i$-th row of the matrix $A$ for $i = 1, \dots, n$.

It follows from the assumption that the sum of elements in each row of $A$ is $0$ that we have

$$A\mathrm{v} = 0.$$

As $\mathrm{v}$ is a nonzero vector, this equality implies that $A$ is a singular matrix.

## Example

To illustrate the proof, let us consider a concrete example.

Let $A = \begin{bmatrix} 1 & -4 & 3 \\ 2 & 5 & -7 \\ 3 & 0 & -3 \end{bmatrix}$.

Then the sum of each row of $A$ is zero.

Now we compute

$$\begin{bmatrix} 1 & -4 & 3 \\ 2 & 5 & -7 \\ 3 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 + (-4) + 3 \\ 2 + 5 + (-7) \\ 3 + 0 + (-3) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This shows that the matrix $A$ is singular because we have $A\mathrm{v} = 0$ for a nonzero vector $\mathrm{v}$.

⑩ Prove that if $n \times n$ matrices $A$ and $B$ are nonsingular, then the product $AB$ is also a nonsingular matrix. (*The Ohio State University*)

### Definition (Nonsingular Matrix)

An $n \times n$ matrix is called nonsingular if the only solution $\mathbf{x} \in \mathbb{R}^n$ of the equation $A\mathbf{x} = \mathbf{x}$ is $\mathbf{x} = 0$.

### Proof.

We give two proofs. The first one uses a property of the determinants of matrices, and the second one uses the definition of nonsingular matrices.

#### Proof 1. (Using Determinant)

Recall that a matrix is nonsingular if and only if its determinant is not zero.

Since $A$ and $B$ are nonsingular, we know that

$$\det(A) \neq 0 \text{ and } \det(B) \neq 0.$$

Then we have using the multiplicative property of the determinant

$$\det(AB) = \det(A)\det(B) \neq 0.$$

Since the determinant of the product $AB$ is not zero, we conclude that $AB$ is a nonsingular matrix.

#### Proof 2. (Using Definition of Nonsingular Matrices)

Suppose that $A, B$ are nonsingular matrices.
This means that if $A\mathbf{x} = 0$ for some the vector $\mathbf{x} \in \mathbb{R}^n$, then we must have $\mathbf{x} = 0$.
Same for $B$.

Suppose that we have $(AB)\mathbf{x} = 0$ for some vector $\mathbf{x} \in \mathbb{R}^n$.
Let $\mathbf{v} = B\mathbf{x} \in \mathbb{R}^n$. Then we have

$$A\mathbf{v} = AB\mathbf{x} = 0.$$

So since $A$ is a nonsingular matrix, we have $\mathbf{v} = 0$, namely, $B\mathbf{x} = 0$.

Since $B$ is nonsingular, this further implies that $\mathbf{x} = 0$.

In summary, whenever $(AB)\mathbf{x} = 0$, we have $\mathbf{x} = 0$.
Therefore, the matrix $AB$ is nonsingular.

⑪ **(a)** Let $A = (a_{ij})$ be an $n \times n$ matrix. Suppose that the entries of the matrix $A$ satisfy the following relation.

$$|a_{ii}| > |a_{i1}| + \cdots + |a_{i\,i-1}| + |a_{i\,i+1}| + \cdots + |a_{in}|$$

for all $1 \leq i \leq n$. Show that the matrix $A$ is nonsingular.
**(b)** Let $B = (b_{ij})$ be an $n \times n$ matrix whose entries satisfy the relation

$$|b_{i\,i}| = 1 \text{ and } |b_{ij}| < \frac{1}{n-1}$$

for all $i$ and $j$ with $i \neq j$. Prove that the matrix $B$ is nonsingular.
**(c)** Determine whether the following matrix is nonsingular or not.

$$C=\begin{bmatrix} \pi & e & e^2/2\pi^2 \\ e^2/2\pi^2 & \pi & e \\ e & e^2/2\pi^2 & \pi \end{bmatrix},$$

where $\pi = 3.14159\dots$, and $e = 2.71828\dots$ is Euler's number (or Napier's constant).

## Hint.

⑫ For (a), assume that the matrix $A$ is nonsingular.
Then the system of linear equation $A\mathbf{x} = 0$ has nonzero solution.

⑬ For (b), apply (a).

⑭ For(c), use part (a). Note $e^2/2\pi^2 = 0.37433\dots$.

## Proof.

### (a) Show that the matrix $A$ is nonsingular.

Suppose that the matrix $A$ is nonsingular. Then the system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \vdots \qquad \vdots \quad &\vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

has nontrivial solution $(x_1,\dots,x_n) \neq (0,\dots,0)$.
Since not all of $x_i$ are zero, let $x_k$ be the nonzero number such that $|x_k|$ is the largest among $|x_i|$.

That is,

$$|x_k| = \max_{1\leq i\leq n} |x_i| > 0.$$

Then from the $k$-th equation in the system, we have

$$-a_{kk}x_k = a_{k1}x_1 + \cdots + a_{k\,k-1}x_{k-1} + a_{k\,k+1}x_{k+1} + \cdots + a_{kn}x_n.$$

Taking the absolute values of both sides and by the triangle inequality, we have

$$\begin{aligned} |a_{kk}||x_k| &= |a_{k1}x_1 + \cdots + a_{k\,k-1}x_{k-1} + a_{k\,k+1}x_{k+1} + \cdots + a_{kn}x_n| \\ &\leq |a_{k1}||x_1| + \cdots + |a_{k\,k-1}||x_{k-1}| + |a_{k\,k+1}||x_{k+1}| + \cdots + |a_{kn}||x_n|. \end{aligned}$$

Since $x_k$ is nonzero, we divide by $|x_k|$ and we have

$$|a_{kk}| \leq |a_{k1}|\frac{|x_1|}{|x_k|} + \cdots + |a_{k\,k-1}|\frac{|x_{k-1}|}{|x_k|} + |a_{k\,k+1}|\frac{|x_{k+1}|}{|x_k|} + \cdots + |a_{kn}|\frac{|x_n|}{|x_k|}.$$

By the choice of $x_k$, the absolute value of $x_k$ is maximal, hence $|x_i|/|x_k| \leq 1$ for all $i$.

Therefore we have

$$|a_{kk}| \leq |a_{i1}| + \cdots + |a_{i\,i-1}| + |a_{i\,i+1}| + \cdots + |a_{in}|.$$

However this contradicts the given condition. Therefore the matrix $A$ is nonsingular.

### (b) Prove that the matrix $B$ is nonsingular.

Note that

$$\begin{aligned} &|b_{i1}| + \cdots + |b_{i\,i-1}| + |b_{i\,i+1}| + \cdots + |b_{in}| \\ &< \frac{1}{n-1} + \cdots + \frac{1}{n-1} = \frac{n-1}{n-1} = 1 = |b_{ii}|. \end{aligned}$$

Therefore the matrix $B$ satisfy the condition of part (a). Therefore from the result of (a), the matrix $B$ is nonsingular.

### (c) Determine whether the matrix is nonsingular or not.

Let us apply part (a) to the matrix $C$.
We have

$$\frac{e^2}{2\pi^2} = 0.37433\dots.$$

Thus we calculate

$$e + \frac{e^2}{2\pi^2} = 3.09262\cdots < \pi = 3.14159\dots$$

Thus the sum of the (absolute value of) off-diagonal entries are less than the diagonal entry. Hence by part (a), the matrix $C$ is nonsingular.

## Comment.

These are simple and interesting criteria for the non singularity of a matrix.
You may easily cook up matrices satisfying (or not satisfying) the conditions.
Why don't you make one and give it to your friend for a math quiz?

⑮ Let $A$ be an $n \times n$ matrix. Suppose that $\mathbf{y}$ is a nonzero row vector such that $\mathbf{y}A = \mathbf{y}$. (Here a row vector means a $1 \times n$ matrix.) Prove that there is a nonzero column vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{x}$. (Here a column vector means an $n \times 1$ matrix.)

## Proof 1.(Without the theory of eigenvalues)

Let $I$ be the $n \times n$ identity matrix. Then we have

$$0_{1\times n} = \mathrm{y}A - \mathrm{y} = \mathrm{y}(A-I),$$

where $0_{1\times n}$ is the row zero vector.

Taking the transpose, we have

$$\begin{aligned} 0_{n\times 1} &= 0_{1\times n}^{\mathrm{T}} = \left(\mathrm{y}(A-I)\right)^{\mathrm{T}} \\ &= (A-I)^{\mathrm{T}}\mathrm{y}^{\mathrm{T}}. \end{aligned}$$

Since the vector $\mathrm{y}$ is nonzero, the transpose $\mathrm{y}^{\mathrm{T}}$ is a nonzero column vector.
Thus, the above equality yields that the matrix $(A-I)^{\mathrm{T}}$ is singular.
It follows that the matrix $A-I$ is singular as well.
Hence there exists a nonzero column vector $\mathrm{x}$ such that

$$(A-I)\mathrm{x} = 0_{n\times 1},$$

and consequently we have

$$A\mathrm{x} = \mathrm{x}$$

for a nonzero column vector $\mathrm{x}$.

## Proof 2. (Using the theory of eigenvalues)

Taking the conjugate of the both sides of the identity $\mathrm{y}A = \mathrm{y}$ , we obtain

$$A^{\mathrm{T}}\mathrm{y}^{\mathrm{T}} = \mathrm{y}^{\mathrm{T}}.$$

Since $\mathrm{y}$ is a nonzero row vector, $\mathrm{y}^{\mathrm{T}}$ is a nonzero column vector.
It follows that $1$ is an eigenvalue of the matrix $A^{\mathrm{T}}$ and $\mathrm{y}^{\mathrm{T}}$ is a corresponding eigenvector.

Since the matrices $A$ and $A^{\mathrm{T}}$ has the same eigenvalues , we deduce that the matrix $A$ has an eigenvalue $1$.
(See part (b) of the post "Transpose of a matrix and eigenvalues and related questions.".)
Let $\mathrm{x}$ be an eigenvector corresponding to the eigenvalue $1$ (by definition $\mathrm{x}$ is nonzero). Then we have

$$A\mathrm{x} = \mathrm{x}$$

as required.

16 Without using properties of determinants
Determine whether each of the following statement is True or False.
**(a)** Suppose that $A$ and $B$ are nonsingular $n \times n$ matrices. Then $A+B$ is nonsingular.
**(b)** If a square matrix has no zero rows or columns, then it has an inverse matrix.
**(c)** Let $A$ be an $m \times n$ matrix.
If the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} \in \mathbb{R}^n$, then the columns of $A$ are linearly independent.
**(d)** Let $A$ be an $m \times n$ matrix.
If the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} \in \mathbb{R}^n$, then the rows of $A$ are linearly independent.
**(e)** The row echelon form of an $3 \times 3$ matrix is invertible.
**(f)** There is a non-zero nonsingular matrix $A$ such that $A^2 = O$.
**(g)** If $A$ and $B$ are invertible $n \times n$ matrices, then $AB = BA$.
**(h)** If $A$ and $B$ are $n \times n$ nonsingular matrices such that $A^2 = I$ and $B^2 = I$, then $(AB)^{-1} = BA$.
**(i)** If $A$ is an $m \times n$ matrix such that $A\mathbf{x} = \mathbf{0}$ for every vector $\mathbf{x}$ in $\mathbb{R}^n$, then $A$ is the $m \times n$ zero matrix.

**(j)** Let $A$ be a $2 \times 2$ nonsingular matrix and let $\mathbf{v}_1$ and $\mathbf{v}_2$ be linearly independent vectors in $\mathbb{R}^2$. Then the vectors $A\mathbf{v}_1, A\mathbf{v}_2$ are linearly independent vectors in $\mathbb{R}^2$.

**17** Let $A$ be an $n \times n$ nonsingular matrix. Prove that the transpose matrix $A^{\mathrm{T}}$ is also nonsingular.

### Definition (Nonsingular Matrix).

By definition, $A^{\mathrm{T}}$ is a **nonsingular matrix** if the only solution to

$$A^{\mathrm{T}}\mathbf{x} = 0$$

is the zero vector $\mathbf{x} = 0$ in $\mathbb{R}^n$.

### Proof.

Suppose that

$$A^{\mathrm{T}}\mathbf{v} = 0 \tag{*}$$

for an $n$-dimensional vector $\mathbf{v}$.

We prove that $\mathbf{v} = 0$.

---

Since $A$ is nonsingular, there exists a vector $\mathbf{u} \in \mathbb{R}^n$ such that

$$A\mathbf{u} = \mathbf{v}.$$

(For those who know the inverse matrix, the vector $\mathbf{u}$ is given by $\mathbf{u} = A^{-1}\mathbf{v}$. The fact can be proved without using the inverse matrix, though.)

Hence we obtain from (*)

$$A^{\mathrm{T}}A\mathbf{u} = 0.$$

It follows that

$$\mathbf{u}^{\mathrm{T}}A^{\mathrm{T}}A\mathbf{u} = \mathbf{u}^{\mathrm{T}}0 = 0.$$

By the property of the transpose, we have

$$\mathbf{u}^{\mathrm{T}}A^{\mathrm{T}} = (A\mathbf{u})^{\mathrm{T}}.$$

Thus we have

$$0 = (A\mathbf{u})^{\mathrm{T}}A\mathbf{u} = \|A\mathbf{u}\|^2.$$

This yields the length $\|A\mathbf{u}\| = 0$, and hence $A\mathbf{u} = 0$.

Since $\mathbf{v} = A\mathbf{u}$, we conclude that $\mathbf{v} = 0$.

Therefore if $A^{\mathrm{T}}\mathrm{v}=0$, then $\mathrm{v}=0$.

This shows that the transpose $A^{\mathrm{T}}$ is nonsingular.

18 Let $A$ be an $n \times (n-1)$ matrix and let $\mathbf{b}$ be an $(n-1)$-dimensional vector. Then the product $A\mathbf{b}$ is an $n$-dimensional vector. Set the $n \times n$ matrix $B=[A_1, A_2, \dots, A_{n-1}, A\mathbf{b}]$, where $A_i$ is the $i$-th column vector of $A$. Prove that $B$ is a singular matrix for any choice of $\mathbf{b}$.

## Definition/Hint.

An $n \times n$ matrix $B$ is **nonsingular** if $A\mathrm{x}=0, \mathrm{x} \in \mathbb{R}^n$ implies that $\mathbf{x}=0$.

Otherwise, the matrix $B$ is called **singular**.

Namely, the matrix $B$ is singular if there exists a nonzero vector $\mathrm{v}$ such that $B\mathrm{v}=0$.

You may use the fact that a matrix is nonsingular if and only if its column vectors are lienarly independent.

We give two proofs. The first one uses the definition of a singular matrix. The second one uses the fact mentioned above.

## Proof (Using Defition).

Let

$$\mathrm{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \end{bmatrix}.$$

Then we have

$$A\mathrm{b}=b_1A_1+b_2A_2+\cdots+b_{n-1}A_{n-1} \tag{*}$$

using the column vectors $A_i$ of $A$.

Let us define the $n$-dimensional vector $\mathrm{v}$ to be

$$\mathrm{v}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ -1 \end{bmatrix}.$$

Since the last entry of $\mathrm{v}$ is $-1$, the vector $\mathrm{v}$ is nonzero.

We calculate

$$
\begin{aligned}
B\mathbf{v} &= [A_1, A_2, \dots, A_{n-1}, A\mathbf{b}] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ -1 \end{bmatrix} \\
&= b_1 A_1 + b_2 A_2 + \cdots + b_{n-1} A_{n-1} + (-1) A\mathbf{b} \\
&\overset{(*)}{=} b_1 A_1 + b_2 A_2 + \cdots + b_{n-1} A_{n-1} - (b_1 A_1 + b_2 A_2 + \cdots + b_{n-1} A_{n-1}) \\
&= \mathbf{0},
\end{aligned}
$$

where $\mathbf{0}$ is the $n$-dimensional zero vector.

Since we have $B\mathbf{v} = \mathbf{0}$ for a nonzero vector $\mathbf{v}$, we conclude that the matrix $B$ is singular.

## Proof (Using the Fact).

Note that a square matrix is singular if and only if its columns vectors are linearly dependent.
We prove that the column vectors $A_1, A_2, \cdots, A_{n-1}, A\mathbf{b}$ of the matrix $B$ are linearly dependent.

Let

$$
\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \end{bmatrix}.
$$

Then we have

$$
A\mathbf{b} = b_1 A_1 + b_2 A_2 + \cdots + b_{n-1} A_{n-1}.
$$

Thus, we have the linear combination of column vectors of $B$

$$
b_1 A_1 + b_2 A_2 + \cdots + b_{n-1} A_{n-1} - A\mathbf{b} = \mathbf{0}.
$$

Since the coefficient in front of the vector $A\mathbf{b}$ is $-1$, the left hand side is a nontrivial linear combination of column vectors of $B$.
This implies that the column vectors of $B$ are linearly dependent, hence the matrix $B$ is singular.

19 An $n \times n$ matrix $A$ is called **nonsingular** if the only vector $\mathbf{x} \in \mathbb{R}^n$ satisfying the equation $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.
Using the definition of a nonsingular matrix, prove the following statements.
**(a)** If $A$ and $B$ are $n \times n$ nonsingular matrix, then the product $AB$ is also nonsingular.
**(b)** Let $A$ and $B$ be $n \times n$ matrices and suppose that the product $AB$ is nonsingular. Then:
1. The matrix $B$ is nonsingular.
2. The matrix $A$ is nonsingular. (You may use the fact that a nonsingular matrix is invertible.)

## Proof.

(a) If $A$ and $B$ are $n \times n$ nonsingular matrix, then the product $AB$ is also nonsingular.

Let $\mathbf{v} \in \mathbb{R}^n$ and suppose that $(AB)\mathbf{x} = 0$.

Our goal is to prove that $\mathbf{x} = 0$.

Let $\mathbf{y} := B\mathbf{x} \in \mathbb{R}^n$. Then we have

$$A\mathbf{y} = A(B\mathbf{x}) = (AB)\mathbf{x} = 0.$$

Since $A$ is nonsingular, this implies that the vector $\mathbf{y} = 0$.

Hence we have $\mathbf{y} = B\mathbf{x} = 0$.

Since $B$ is nonsingular, this further implies that $\mathbf{x} = 0$.

It follows that if $(AB)\mathbf{x} = 0$, then we must have $\mathbf{x} = 0$.

By definition, this means that the matrix $AB$ is nonsingular.

### (b)-1. If $AB$ is nonsingular, then $B$ is nonsingular.

Suppose that $B\mathbf{x} = 0$. We prove that $\mathbf{x} = 0$.

Since $B\mathbf{x} = 0$, it yields that

$$(AB)\mathbf{x} = A(B\mathbf{x}) = A0 = 0.$$

As the matrix $AB$ is nonsingular, it follows from $(AB)\mathbf{x} = 0$ that $\mathbf{x} = 0$.

This proves that the matrix $B$ is nonsingular.

### (b)-2. If $AB$ is nonsingular, then $A$ is nonsingular.

By part (1), we know that $B$ is nonsingular, hence it is invertible.

The inverse matrix $B^{-1}$ and the matrix $AB$ are both nonsingular.

Hence it follows from part (a) that the product of $AB$ and $B^{-1}$ is also nonsingular.

Thus,

$$A = (AB)B^{-1}$$

is a nonsingular matrix.

## Nonsingular if and only if Invertible

For the proof of the fact we used in the proof of (b)-2 that a matrix is nonsingular if and only if it is invertible, see the post↴

A Matrix is Invertible If and Only If It is Nonsingular

20 Let $A$ be a singular $n \times n$ matrix.
Let

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

be unit vectors in $\mathbb{R}^n$.

Prove that at least one of the following matrix equations

$$A\mathbf{x} = \mathbf{e}_i$$

for $i = 1, 2, \dots, n$, must have no solution $\mathbf{x} \in \mathbb{R}^n$.

### Proof.

Assume on the contrary that each matrix equation $A\mathrm{x} = \mathrm{e}_i$ has a solution.
Let $\mathrm{b}_i \in \mathbb{R}^n$ be a solution of $A\mathrm{x} = \mathrm{e}_i$ for each $i = 1, \dots, n$.
That is, we have

$$A\mathrm{b}_i = \mathrm{e}_i.$$

Let $B = [\mathrm{b}_1, \mathrm{b}_2, \dots, \mathrm{b}_n]$ be the $n \times n$ matrix whose $i$-th column vector is $\mathrm{b}_i$.

Then we have

$$\begin{aligned} AB &= A[\mathrm{b}_1, \mathrm{b}_2, \dots, \mathrm{b}_n] \\ &= [A\mathrm{b}_1, A\mathrm{b}_2, \dots, A\mathrm{b}_n] \\ &= [\mathrm{e}_1, \mathrm{e}_2, \dots, \mathrm{e}_n] = I, \end{aligned}$$

where $I$ is the $n \times n$ identity matrix.

Since $I$ is the nonsingular matrix, the matrix $A$ must also be nonsingular.
However this contradicts the assumption that $A$ is singular.
It follows that at least one of the matrix equations $A\mathrm{x} = \mathrm{e}_i$ has no solution.

**21** Using the numbers appearing in

$$\pi = 3.1415926535897932384626433832795028841971693993751058209749\dots$$

we construct the matrix

$$A = \begin{bmatrix} 3 & 14 & 1592 & 65358 \\ 97932 & 38462643 & 38 & 32 \\ 7950 & 2 & 8841 & 9716 \\ 939937510 & 5820 & 974 & 9 \end{bmatrix}.$$

Prove that the matrix $A$ is nonsingular.

### Proof.

To show that the matrix $A$ is nonsingular, it suffices to prove that $\det(A) \neq 0$.

One way is to compute the determinant of $A$ directly.

However, as the numbers in $A$ are quite large for hand computation, the direct calculation must be tedious.

So we consider an alternative method.

Note that we do not have to find the exact value of $\det(A)$, but we just need to know $\det(A) \neq 0$.

Thus, it suffices to show that $\det(A)$ is odd. ($0$ is an even number.)

This suggests that considering the matrix modulo $2$ is helpful.

---

Let $\bar{A}$ be the matrix whose $(i,j)$-entry is the $(i,j)$-entry of $A$ modulo $2$.

That is,

$$\bar{A} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(Remark that the diagonal entries of $A$ are odd, and off-diagonal entries are even.)

Since $\det(A)$ is a polynomial of entries of $A$, we have

$$\begin{aligned} \det(A) &\equiv \det(\bar{A}) \pmod{2} \\ &= 1. \end{aligned}$$

It follows that $\det(A)$ is odd, and in particular $\det(A) \neq 0$.

Thus the matrix $A$ is nonsingular.

## What's $\det(A)$ anyway?

Just for the record, the determinant of $A$ is

$$\det(A) = -20330769121541702776233175.$$

## Beautiful Formulas for $\pi$

This problem was nothing to do with the number $\pi$ (except we used the digits of $\pi$) and the matrix is far from beautiful.

(Although the method we used is beautiful.)

Check out the post

Beautiful formulas for pi=3.14…

for beautiful formulas containing $\pi$ like

$$\pi = \cfrac{4}{1 + \cfrac{1^2}{2 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cfrac{7^2}{2 + \cdots}}}}}$$

# Inverse Matrices

**Definition**

① An $n \times n$ matrix $A$ is said to be **invertible** if there exists an $n \times n$ matrix $B$ such that $AB = BA = I$. Such a matrix $B$ is unique and called the **inverse matrix** of $A$, denoted by $A^{-1}$

**Summary**

Let $A, B$ be $n \times n$ matrices.

① $A$ is invertible if and only if $\mathrm{rref}([A \mid I_n]) = [I_n \mid A']$ for some $n \times n$ matrix $A'$. In this case, $A' = A^{-1}$.

② $A$ is invertible if and only if $A$ is nonsingular.

③ If $A, B$ are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$

④ A $2 \times 2$ matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if the determinant $\det(A) = ad - bc \neq 0$.

If $A$ is invertible, then the inverse matrix is given by $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

⑤ If $A$ is invertible, then $A^{\mathrm{T}}$ is invertible and $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$.

## Problems

① For each of the following $3 \times 3$ matrices $A$, determine whether $A$ is invertible and find the inverse $A^{-1}$ if exists by computing the augmented matrix $[A|I]$, where $I$ is the $3 \times 3$ identity matrix.

**(a)** $A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{bmatrix}$

**(b)** $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -3 & 2 \\ 3 & 6 & -2 \end{bmatrix}$.

### Elementary Row Operations and Inverse Matrices

Recall the following procedure of testing the invertibility of $A$ as well as finding the inverse matrix if exists.

If the augmented matrix $[A|I]$ is transformed into a matrix of the form $[I|B]$, then the matrix $A$ is invertible and the inverse matrix $A^{-1}$ is given by $B$.

If the reduced row echelon form matrix for $[A|I]$ is not of the form $[I|B]$, then the matrix $A$ is not invertible.

## Solution.

(a) $A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{bmatrix}$

We apply the elementary row operations as follows.

We have

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_2-2R_1} \left[\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & -3 & 4 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2\leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & -3 & 4 & -2 & 1 & 0 \end{array}\right] \xrightarrow[R_3+3R_2]{R_1-3R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & -3 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 3 \end{array}\right]$$

$$\xrightarrow[R_2+R_3]{R_1-R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & -6 \\ 0 & 1 & 0 & -2 & 1 & 4 \\ 0 & 0 & 1 & -2 & 1 & 3 \end{array}\right].$$

The left $3 \times 3$ part of the last matrix is the identity matrix.

This implies that $A$ is invertible and the inverse matrix is given by the right $3 \times 3$ matrix.

Hence

$$A^{-1} = \begin{bmatrix} 3 & -1 & -6 \\ -2 & 1 & 4 \\ -2 & 1 & 3 \end{bmatrix}.$$

(b) $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -3 & 2 \\ 3 & 6 & -2 \end{bmatrix}.$

Now we consider the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -3 & 2 \\ 3 & 6 & -2 \end{bmatrix}.$

Applying elementary row operations, we obtain

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ -1 & -3 & 2 & 0 & 1 & 0 \\ 3 & 6 & -2 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3-3R_1]{R_2+R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -3 & 4 & 1 & 1 & 0 \\ 0 & 6 & -8 & -3 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -3 & 4 & 1 & 1 & 0 \\ 0 & 0 & 0 & -5 & -2 & 1 \end{array}\right] \xrightarrow{\frac{-1}{3}R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -4/3 & -1/3 & -1/3 & 0 \\ 0 & 0 & 0 & -5 & -2 & 1 \end{array}\right].$$

The last matrix is in reduced row echelon form but the left $3 \times 3$ part is not the identity matrix $I$.

It follows that the matrix $A$ is not invertible.

## Related Question.

**Problem.**

Find the inverse matrix of

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

if it exists. If you think there is no inverse matrix of $A$, then give a reason.

② Let A be the matrix

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Is the matrix $A$ invertible? If not, then explain why it isn't invertible. If so, then find the inverse.
(*The Ohio State University*)

## Hint.

Consider the augmented matrix $[A|I]$, where $I$ is the $3 \times 3$ identity matrix.

## Solution.

Consider the augmented matrix $[A|I]$ with the $3 \times 3$ identity matrix $I$.

Reduce $[A|I]$ using elementary row operations as follows.

$$\begin{aligned} [A|I] &= \left[\begin{array}{rrr|rrr} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \\ &\xrightarrow{R_2+R_3} \left[\begin{array}{rrr|rrr} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \\ &\xrightarrow{R_1+R_2} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]. \end{aligned}$$

This is in reduced row echelon form and the left $3 \times 3$ part is the identity matrix. Hence $A$ is invertible and the inverse matrix is

$$A^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

③ Find the inverse matrix of

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

if it exists. If you think there is no inverse matrix of $A$, then give a reason.
(*The Ohio State University*)

## Solution.

To check whether the matrix $A$ has the inverse matrix and to find the inverse matrix if exist at once, we consider the augmented matrix $[A \mid I]$, where $I$ is the $3 \times 3$ identity matrix.
We apply the elementary row operations as follows.

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_3 - R_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{-R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & -1 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{R_1 - R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right] \xrightarrow[R_2 - R_3]{R_1 - R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right].$$

The left $3 \times 3$ matrix part became the identity matrix $I$, thus $A$ is invertible (since it is row equivalent to $I$ ), and the inverse matrix $A^{-1}$ is given by the right $3 \times 3$ matrix. Thus we have

$$A^{-1} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

### Double check

Once you obtained the inverse matrix $A^{-1}$, then you should check that whether $A^{-1}A$ is the identity matrix.
If not, it means that you did a computational mistake somewhere.

④ Let $A$ be the following $3 \times 3$ upper triangular matrix.

$$A = \begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix},$$

where $x, y, z$ are some real numbers. Determine whether the matrix $A$ is invertible or not. If it is invertible, then find the inverse matrix $A^{-1}$.

## Solution.

We form the augmented matrix

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & x & y & 1 & 0 & 0 \\ 0 & 1 & z & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

and apply elementary row operations as follows.

$$\left[\begin{array}{ccc|ccc} 1 & x & y & 1 & 0 & 0 \\ 0 & 1 & z & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1 - xR_2} \left[\begin{array}{ccc|ccc} 1 & 0 & y - xz & 1 & -x & 0 \\ 0 & 1 & z & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_2 - zR_3]{R_1 - (y - xz)R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -x & xz - y \\ 0 & 1 & 0 & 0 & 1 & -z \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

We could reduce the matrix $A$ into the identity matrix $I$.

Thus, the matrix $A$ is invertible and the right $3 \times 3$ matrix is the inverse matrix of $A^{-1}$.

Hence,

$$A^{-1} = \begin{bmatrix} 1 & -x & xz - y \\ 0 & 1 & -z \\ 0 & 0 & 1 \end{bmatrix}.$$

5 For which choice(s) of the constant $k$ is the following matrix invertible?

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & k \\ 1 & 4 & k^2 \end{bmatrix}.$$

(*Johns Hopkins University*)

## Hint.

An $n \times n$ matrix is invertible if and only if its rank is $n$.

The rank of a matrix is the number of nonzero rows of a (reduced) row echelon form matrix that is row equivalent to the given matrix.

## Solution.

We compute the rank of the matrix $A$.

Applying elementary row operations, we obtain

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & k \\ 1 & 4 & k^2 \end{bmatrix} \xrightarrow[R_3 - R_1]{R_2 - R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & k - 1 \\ 0 & 3 & k^2 - 1 \end{bmatrix} \xrightarrow[R_3 - 3R_2]{R_1 - R_2} \begin{bmatrix} 1 & 0 & 2 - k \\ 0 & 1 & k - 1 \\ 0 & 0 & k^2 - 3k + 2 \end{bmatrix}.$$

The last matrix is in row echelon form.

Note that $A$ is an invertible matrix if and only if its rank is $3$.

Therefore the $(3, 3)$-entry of the last matrix must be nonzero: $k^2 - 3k + 2 = (k - 1)(k - 2) \neq 0$.

It follows that the matrix $A$ is invertible for any $k$ except $k = 1, 2$.

⑥ Suppose that $M, P$ are two $n \times n$ non-singular matrix. Prove that there is a matrix $N$ such that $MN = P$.

### Proof.

As non-singularity and invertibility are equivalent, we know that $M$ has the inverse matrix $M^{-1}$.

Let us think backwards. Suppose that we have $MN = P$ for some matrix $N$, which we want to find. Then multiply $M^{-1}$ on the left of the equation $MN = P$ yields $N = M^{-1}P$.

This is the matrix we are looking for, as $MN = M(M^{-1}P) = P$.

⑦ Let $A$ be an $n \times n$ matrix satisfying $A^2 + c_1 A + c_0 I = O$, where $c_0, c_1$ are scalars, $I$ is the $n \times n$ identity matrix, and $O$ is the $n \times n$ zero matrix. Prove that if $c_0 \neq 0$, then the matrix $A$ is invertible (nonsingular). How about the converse? Namely, is it true that if $c_0 = 0$, then the matrix $A$ is not invertible?

### Proof.

#### If $c_0 \neq 0$, then the matrix $A$ is invertible

Suppose first that $c_0 \neq 0$.

Then we have

$$\begin{aligned} A^2 + c_1 A &= -c_0 I \\ \Leftrightarrow A(A + c_1 I) &= -c_0 I \\ \Leftrightarrow A\left(\frac{-1}{c_0}(A + c_1 I)\right) &= I. \end{aligned}$$

It is in the last step that we needed to assume $c_0 \neq 0$.

Thus, if we put

$$B = \frac{-1}{c_0}(A + c_1 I),$$

then we have proved that

$$AB = I.$$

Similarly, one can check that $BA = A$. Hence $B$ is the inverse matrix of $A$.

Namely,

$$A^{-1} = \frac{-1}{c_0}(A + c_1 I).$$

This proves that when $c_0 \neq 0$ the matrix $A$ is invertible.

Is it true that if $c_0 = 0$, then the matrix $A$ is not invertible?

Next, let us consider the case $c_0 = 0$.
We claim that the matrix $A$ can be invertible even $c_0 = 0$.

For example, if $A = I$, then $A$ satisfies

$$A^2 - A = O.$$

(Thus, $c_1 = -1$ and $c_0 = 0$.)

Since the identity matrix is invertible, the condition $c_0 = 0$ does not force the matrix $A$ to be non-invertible.

**8** A square matrix $A$ is called **idempotent** if $A^2 = A$. Show that a square invertible idempotent matrix is the identity matrix.

## Proof.

Let $A$ be an $n \times n$ invertible idempotent matrix.

Since $A$ is invertible, the inverse matrix $A^{-1}$ of $A$ exists and it satisfies $A^{-1}A = I_n$, where $I_n$ is the $n \times n$ identity matrix.

Since $A$ is idempotent, we have $A^2 = A$.
Multiplying this equality by $A^{-1}$ from the left, we get $A^{-1}A^2 = A^{-1}A$. Using the fact that $A^{-1}A = I_n$, we obtain $A = I_n$.

The proof is completed.

## Related Question.

Give it a try with the following problems about idempotent matrices.

**Problem**.
**(a)** Let $\mathbf{u}$ be a vector in $\mathbb{R}^n$ with length $1$.
Define the matrix $P$ to be $P = \mathbf{u}\mathbf{u}^{\mathrm{T}}$.

Prove that $P$ is an idempotent matrix.

---

**(b)** Suppose that $\mathbf{u}$ and $\mathbf{v}$ be unit vectors in $\mathbb{R}^n$ such that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.
Let $Q = \mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}$.

Prove that $Q$ is an idempotent matrix.

---

**(c)** Prove that each nonzero vector of the form $a\mathbf{u} + b\mathbf{v}$ for some $a, b \in \mathbb{R}$ is an eigenvector corresponding to the eigenvalue $1$ for the matrix $Q$ in part (b).

9

Find the inverse matrix of $A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{bmatrix}$ if it exists. If you think there is no inverse matrix of $A$, then give a reason.

---

### (a) Find the inverse matrix if exists.

We consider the augmented matrix $[A \mid I]$, where $I$ is the $3 \times 3$ identity matrix, and apply elementary row operations as follows. We have

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3-2R_1]{R_2-R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 1 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 0 \end{array}\right] \xrightarrow{-R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right]$$

$$\xrightarrow[R_2+R_3]{R_1-R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right].$$

Therefore, we could reduce the matrix $A$ into the identity matrix ($A$ is row equivalent to $I$) and thus $A$ is invertible, and the inverse matrix is given by the right half of the last matrix.

Hence the inverse matrix is

$$A^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

### Find a nonsingular $2 \times 2$ matrix $A$ such that $A^3 = A^2B - 3A^2$.

Suppose that $A$ is a nonsingular matrix such that $A^3 = A^2B - 3A^2$.
Since $A$ is nonsingular, it is invertible and thus the inverse matrix $A^{-1}$ exists.

Then we have

$$\begin{aligned} A &= A^{-2}A^3 \\ &= A^{-2}(A^2B - 3A^2) \\ &= A^{-2}A^2B - 3A^{-2}A^2 \\ &= IB - 3I \\ &= B - 3I \\ &= \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}. \end{aligned}$$

To prove that this matrix is nonsingular we calculate the determinant of $A$.
(Recall that if the determinant of a matrix is nonzero, then it is invertible, hence nonsingular.)
The determinant of $2 \times 2$ matrix $A$ is

$$\det(A) = 1 \cdot 3 - 1 \cdot 2 = 1 \neq 0.$$

Since the determinant of $A$ is nonzero, we conclude that $A$ is nonsingular.

Thus, the nonsingular matrix $A$ satisfying $A^3 = A^2B - 3A^2$ must be

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}.$$

10 Find a nonsingular $2 \times 2$ matrix $A$ such that $A^3 = A^2B - 3A^2$, where $B = \begin{bmatrix} 4 & 1 \\ 2 & 6 \end{bmatrix}$. Verify that the matrix $A$ you obtained is actually a nonsingular matrix.

> (a) Find the inverse matrix if exists.

We consider the augmented matrix $[A \mid I]$, where $I$ is the $3 \times 3$ identity matrix, and apply elementary row operations as follows. We have

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{R_2-R_1 \\ R_3-2R_1}} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 1 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 0 \end{array}\right] \xrightarrow{-R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{\substack{R_1-R_3 \\ R_2+R_3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right].$$

Therefore, we could reduce the matrix $A$ into the identity matrix ($A$ is row equivalent to $I$) and thus $A$ is invertible, and the inverse matrix is given by the right half of the last matrix.

Hence the inverse matrix is

$$A^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

> Find a nonsingular $2 \times 2$ matrix $A$ such that $A^3 = A^2B - 3A^2$.

Suppose that $A$ is a nonsingular matrix such that $A^3 = A^2B - 3A^2$.

Since $A$ is nonsingular, it is invertible and thus the inverse matrix $A^{-1}$ exists.

Then we have

$$\begin{aligned} A &= A^{-2}A^3 \\ &= A^{-2}(A^2B - 3A^2) \\ &= A^{-2}A^2B - 3A^{-2}A^2 \\ &= IB - 3I \\ &= B - 3I \\ &= \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}. \end{aligned}$$

To prove that this matrix is nonsingular we calculate the determinant of $A$.

(Recall that if the determinant of a matrix is nonzero, then it is invertible, hence nonsingular.)

The determinant of $2 \times 2$ matrix $A$ is

$$\det(A) = 1 \cdot 3 - 1 \cdot 2 = 1 \neq 0.$$

Since the determinant of $A$ is nonzero, we conclude that $A$ is nonsingular.

Thus, the nonsingular matrix $A$ satisfying $A^3 = A^2B - 3A^2$ must be

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}.$$

---

⑪ Determine whether there exists a nonsingular matrix $A$ if $A^2 = AB + 2A$, where $B$ is the following matrix. If such a nonsingular matrix $A$ exists, find the inverse matrix $A^{-1}$.

**(a)**

$$B = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 1 & 2 & -2 \end{bmatrix}$$

**(b)**

$$B = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 2 & 1 & -4 \end{bmatrix}.$$

---

## Solution.

Suppose that a nonsingular matrix $A$ satisfying $A^2 = AB + 2A$ exists.

Then $A$ is invertible since $A$ is nonsingular, and thus the inverse $A^{-1}$ exists.

Multiplying by $A^{-1}$ on the left, we have

$$\begin{aligned} A &= A^{-1}A^2 = A^{-1}(AB + 2A) \\ &= A^{-1}AB + 2A^{-1}A \\ &= B + 2I, \end{aligned}$$

where $I$ is the $3 \times 3$ identity matrix.

Therefore, if such a nonsingular matrix exists, it must be

$$A = B + 2I. \tag{*}$$

### (a) The first case

Let us consider the case

$$B = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 1 & 2 & -2 \end{bmatrix}.$$

In this case, we have from (*)

$$A = B + 2I = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 2 & 0 \end{bmatrix}.$$

We still need to check that this matrix is in fact a nonsingular matrix.

To check the non-singularity and to find the inverse matrix at once, we consider the augmented matrix $[A \mid I]$ and apply elementary row operations.

We have

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 - R_1} \left[\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array}\right] \xrightarrow[R_3 - R_2]{R_1 - R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 + R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -2 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right].$$

The left part of the last matrix is the identity matrix, and thus the matrix $A$ is invertible and the inverse matrix is the right half:

$$A^{-1} = \begin{bmatrix} 0 & -2 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}.$$

### (b) The second case

Next let us consider the case

$$B = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 2 & 1 & -4 \end{bmatrix}.$$

By (*), if a nonsingular matrix $A$ exists, it must be

$$A = B + 2I = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & 1 & -2 \end{bmatrix}.$$

We need to determine whether this matrix is actually nonsingular.

In fact, we prove that this matrix is singular.

That is, we show that $A\mathbf{x} = 0$ has a nonzero solution.

Consider the augmented matrix $[A \mid 0]$. By Gauss-Jordan elimination, we have

$$[A \mid 0] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 \end{array}\right] \xrightarrow{R_3-2R_1}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right] \xrightarrow[\ ]{\substack{R_1-R_2 \\ R_3+R_2}} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The last matrix is in reduced row echelon form and it has a zero row. From this, we see that $x_3$ must be a free variable, and the matrix $A$ is singular.
(The general solution is $x_1 = x_3, x_2 = 0$. Thus for example, $x_1 = 1, x_2 = 0, x_3 = 1$ is a nonzero solution of $A\mathrm{x} = 0$ .)

Thus, we conclude that there is no nonsingular matrix $A$ satisfying $A^2 = AB + 2A$ in this case.

⑫ Determine whether there exists a nonsingular matrix $A$ if $A^4 = ABA^2 + 2A^3$, where $B$ is the following matrix. $B = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 2 & 1 & -4 \end{bmatrix}$. If such a nonsingular matrix $A$ exists, find the inverse matrix $A^{-1}$.
(*The Ohio State University, Linear Algebra Final Exam Problem*)

## Solution.

Assume that $A$ is nonsingular.
Then the inverse matrix $A^{-1}$ exists and we have

$$\begin{aligned} A &= A^{-1}A^4A^{-2} \\ &= A^{-1}(ABA^2 + 2A^3)A^{-2} \\ &= A^{-1}ABA^2A^{-2} + 2A^{-1}A^3A^{-2} \\ &= B + 2I, \end{aligned}$$

where $I$ is the $3 \times 3$ identity matrix.

Thus we have

$$A = B + 2I = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & 1 & -2 \end{bmatrix}.$$

We compute the determinant of $A$ by the second row cofactor expansion as follows.

$$\det(A) = \begin{vmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & 1 & -2 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 2 & -2 \end{vmatrix} = 0.$$

This contradicts the assumption that $A$ is nonsingular.

Therefore, there is no nonsingular matrix $A$ satisfying the relation $A^4 = ABA^2 + 2A^3$ .

13 Let $A, B, C$ be the following $3 \times 3$ matrices.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 5 \end{bmatrix}, C = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 5 & 6 \\ 3 & 0 & 1 \end{bmatrix}.$$

Then compute and simplify the following expression.

$$(A^{\mathrm{T}} - B)^{\mathrm{T}} + C(B^{-1}C)^{-1}.$$

(*The Ohio State University*)

## Solution.

We first use the properties of transpose matrices and inverse matrices and simplify the expression.
Note that we have

$$(A^{\mathrm{T}} - B)^{\mathrm{T}} = (A^{\mathrm{T}})^{\mathrm{T}} - B^{\mathrm{T}} = A - B$$

since the double transpose $(A^{\mathrm{T}})^{\mathrm{T}} = A$ and $B$ is a symmetric matrix.

Also, note that we have

$$(B^{-1}C)^{-1} = C^{-1}(B^{-1})^{-1} = C^{-1}B$$

since $(B^{-1})^{-1} = B$. Care must be taken when you distribute the inverse sign in the first equality. We needed to switch the order of the product.
Then we have

$$C(B^{-1}C)^{-1} = CC^{-1}B = IB = B,$$

where $I$ is the $3 \times 3$ identity matrix.

Therefore, the given expression can be simplified into

$$(A^{\mathrm{T}} - B)^{\mathrm{T}} + C(B^{-1}C)^{-1} = A - B + B = A.$$

Hence we have

$$(A^{\mathrm{T}} - B)^{\mathrm{T}} + C(B^{-1}C)^{-1} = A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

### Common Mistakes

Here are two common mistakes.

The first one is

$$(B^{-1}C)^{-1} = (B^{-1})^{-1}C^{-1} \quad \text{(This is wrong!!)}.$$

Note that in general, we have

$$(AB)^{-1} = B^{-1}A^{-1} \quad \text{(Note the order of products)}.$$

The second common mistake is that

$$CBC^{-1} = CC^{-1}B = B \quad \text{(The first equality is wrong!!)}.$$

Recall that in general matrix multiplication is not commutative, meaning that

$$AB \neq BA.$$

And surprisingly, if you combine these two mistakes, you get the right answer.

$$C(B^{-1}C)^{-1} = CBC^{-1} = B. \quad \text{(This is wrong!!)}.$$

However, these mistakes show that you didn't understand matrix operations including transpose and inverse matrices.

---

**14** Let $A$ be the coefficient matrix of the system of linear equations

$$\begin{aligned} -x_1 - 2x_2 &= 1 \\ 2x_1 + 3x_2 &= -1. \end{aligned}$$

**(a)** Solve the system by finding the inverse matrix $A^{-1}$.

**(b)** Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be the solution of the system obtained in part (a). Calculate and simplify

$$A^{2017}\mathbf{x}.$$

(*The Ohio State University*)

---

## Solution.

### (a) Solve the system by finding the inverse matrix $A^{-1}$.

The coefficient matrix $A$ of the system is

$$A = \begin{bmatrix} -1 & -2 \\ 2 & 3 \end{bmatrix}.$$

The determinant of the matrix $A$ is $\det(A) = (-1)(3) - (-2)(2) = 1 \neq 0$, and thus $A$ is invertible and the inverse matrix can be found by the $2 \times 2$ inverse matrix formula

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\det(A) = ad - bc \neq 0$.

Thus

$$A^{-1} = \begin{bmatrix} 3 & 2 \\ -2 & -1 \end{bmatrix}.$$

(Or you could form the augmented matrix $[A \mid I]$ and find $A^{-1}$.)

---

Now the given system of linear equations can be written as

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Multiplying by $A^{-1}$ on the left, we obtain

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Thus the solution of the system is $x_1 = 1, x_2 = -1$, or in the vector form

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

### (b) Calculate and simplify $A^{2017}\mathbf{x}$

Since $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is the solution of the system $A\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, we see that

$$A\mathbf{x} = \mathbf{x}.$$

From this, we have

$$\begin{aligned} A^2\mathbf{x} &= AA\mathbf{x} = A\mathbf{x} = \mathbf{x} \\ A^3\mathbf{x} &= AA^2\mathbf{x} = A\mathbf{x} = \mathbf{x} \\ A^4\mathbf{x} &= AA^3\mathbf{x} = A\mathbf{x} = \mathbf{x} \\ &\vdots \end{aligned}$$

Repeating this, we see that

$$A^{2017}\mathbf{x} = \mathbf{x}.$$

(To be more precise, you can prove that $A^n\mathbf{x} = \mathbf{x}$ for any positive integer $n$ by mathematical induction on $n$.)

Thus, we have

$$A^{2017}\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

15 **(a)** Let $A$ be a $6 \times 6$ matrix and suppose that $A$ can be written as $A = BC$, where $B$ is a $6 \times 5$ matrix and $C$ is a $5 \times 6$ matrix. Prove that the matrix $A$ cannot be invertible.
**(b)** Let $A$ be a $2 \times 2$ matrix and suppose that $A$ can be written as $A = BC$, where $B$ is a $2 \times 3$ matrix and $C$ is a $3 \times 2$ matrix. Can the matrix $A$ be invertible?

## Solution.

### (a) Prove that the matrix $A$ cannot be invertible.

Since $C$ is a $5 \times 6$ matrix, the equation

$$C\mathbf{x} = 0$$

has a nonzero solution $\mathbf{x}_1$.

(There are more variables than equations in the system $C\mathbf{x} = 0$.)

It follows that we have

$$A\mathbf{x}_1 = BC\mathbf{x}_1 = B0 = 0.$$

Since the vector $\mathbf{x}_1$ is nonzero, the matrix $A$ is a singular matrix, hence $A$ is not invertible.

### (b) Can the matrix $A$ be invertible?

The answer is yes. For example consider the following $2 \times 3$ matrix $B$ and $3 \times 2$ matrix $C$:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then we have

$$A = BC = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and the matrix $A$ is invertible.

**16** For a real number $a$, consider $2 \times 2$ matrices $A, P, Q$ satisfying the following five conditions.
(1) $A = aP + (a+1)Q$, (2) $P^2 = P$, (3) $Q^2 = Q$, (4) $PQ = O$, (5) $QP = O$, where $O$ is the $2 \times 2$ zero matrix. Then do the following problems.
**(a)** Prove that $(P+Q)A = A$.
**(b)** Suppose $a$ is a positive real number and let $A = \begin{bmatrix} a & 0 \\ 1 & a+1 \end{bmatrix}$. Then find all matrices $P, Q$ satisfying conditions (1)-(5).
**(c)** Let $n$ be an integer greater than $1$. For any integer $k$, $2 \leq k \leq n$, we define the matrix $A_k = \begin{bmatrix} k & 0 \\ 1 & k+1 \end{bmatrix}$.
Then calculate and simplify the matrix product $A_n A_{n-1} A_{n-2} \cdots A_2$.
(*Tokyo University Entrance Exam 2007*)

## Solution.

### (a) Prove that $(P+Q)A = A$ .

We have

$$\begin{aligned}(P+Q)A &\overset{(1)}{=} (P+Q)(aP+(a+1)Q) \\ &= aP^2 + (a+1)PQ + aQP + (a+1)Q^2 \\ &= aP + (a+1)O + aO + (a+1)Q \\ &\quad [\text{ by } (2), (3), (4), (5)] \\ &= aP + (a+1)Q \overset{(1)}{=} A.\end{aligned}$$

Hence, we obtain

$$(P+Q)A = A$$

as required.

### (b) Find all matrices $P, Q$

Note that the matrix $A = \begin{bmatrix} a & 0 \\ 1 & a+1 \end{bmatrix}$ is invertible since the determinant

$$\det(A) = \begin{vmatrix} a & 0 \\ 1 & a+1 \end{vmatrix} = a(a+1) \neq 0.$$

By part (a), we know $(P+Q)A = A$ .
Multiplying this by $A^{-1}$ from the right, we have

$$P + Q = I,$$

where $I$ is the $2 \times 2$ identity matrix.
Substituting $Q = I - P$ into the equality $A = aP + (a+1)Q$ of (1), we have

$$\begin{aligned} A &= aP + (a+1)(I-P) \\ &= (a+1)I - P. \end{aligned}$$

Thus, we have

$$\begin{aligned} P &= (a+1)I - A \\ &= \begin{bmatrix} a+1 & 0 \\ 0 & a+1 \end{bmatrix} - \begin{bmatrix} a & 0 \\ 1 & a+1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \end{aligned}$$

and thus

$$\begin{aligned} Q &= I - P \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

It is straightforward to check that the matrices

$$P = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

satisfy conditions (1)-(5).
Hence these are the only matrices satisfying conditions (1)-(5).

### (c) Calculate and simplify the matrix product $A_n A_{n-1} A_{n-2} \cdots A_2$

In part (2), we showed that for any positive integer $k$

$$A_k = kP + (k+1)Q,$$

where

$$P = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

(We just applied the result of (b) with $a = k$.)

Thus we have

$$\begin{aligned} &A_n A_{n-1} \cdots A_2 \\ &= (nP + (n+1)Q)\,((n-1)P + nQ) \cdots (2P + 3Q) \\ &= n!P + \frac{(n+1)!}{2} Q \end{aligned}$$

We used conditions (4) and (5) in the second equality.

Using the explicit matrices for $P$ and $Q$, we have

$$\begin{aligned} &n!P + \frac{(n+1)!}{2} Q \\ &= n! \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} + \frac{(n+1)!}{2} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} n! & 0 \\ -n! + \frac{(n+1)!}{2} & \frac{(n+1)!}{2} \end{bmatrix} \\ &= \begin{bmatrix} n! & 0 \\ n!\frac{n-1}{2} & \frac{(n+1)!}{2} \end{bmatrix}. \end{aligned}$$

Note that in the last step, we computed

$$\begin{aligned} &-n! + \frac{(n+1)!}{2} = -n! + \frac{(n+1)n!}{2} \\ &= n!(-1 + \frac{n+1}{2}) \\ &= n!\frac{n-1}{2}. \end{aligned}$$

In conclusion, we have obtained

$$A_n A_{n-1} \cdots A_2 = \begin{bmatrix} n! & 0 \\ n!\frac{n-1}{2} & \frac{(n+1)!}{2} \end{bmatrix}.$$

## Comment.

Another way to solve (c) is that one first guesses the formula we obtained and prove it by mathematical induction.

⑰ Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^n$. Then the dot product $\mathbf{v} \cdot \mathbf{v} = \mathbf{v}^{\mathrm{T}} \mathbf{v} \neq 0$. Set $a := \frac{2}{\mathbf{v}^{\mathrm{T}} \mathbf{v}}$ and define the $n \times n$ matrix $A$ by $A = I - a\mathbf{v}\mathbf{v}^{\mathrm{T}}$, where $I$ is the $n \times n$ identity matrix. Prove that $A$ is a symmetric matrix and $AA = I$. Conclude that the inverse matrix is $A^{-1} = A$.

## Proof.

### $A$ is symmetric

We first show that the matrix $A$ is symmetric.

We calculate using properties of transpose

$$\begin{aligned} A^{\mathrm{T}} &= (I - a\mathrm{v}\mathrm{v}^{\mathrm{T}})^{\mathrm{T}} && \text{by definition of } A \\ &= I^{\mathrm{T}} - (a\mathrm{v}\mathrm{v}^{\mathrm{T}})^{\mathrm{T}} \\ &= I - a(\mathrm{v}^{\mathrm{T}})^{\mathrm{T}}\mathrm{v}^{\mathrm{T}} \\ &= I - a\mathrm{v}\mathrm{v}^{\mathrm{T}} \\ &= A && \text{by definition of } A. \end{aligned}$$

Hence we have $A^{\mathrm{T}} = A$, and thus $A$ is symmetric.

### $AA = I$ and $A^{-1} = A$

Next, we prove that $AA = I$.

We compute

$$\begin{aligned} AA &= (I - a\mathrm{v}\mathrm{v}^{\mathrm{T}})(I - a\mathrm{v}\mathrm{v}^{\mathrm{T}}) \\ &= I(I - a\mathrm{v}\mathrm{v}^{\mathrm{T}}) - a\mathrm{v}\mathrm{v}^{\mathrm{T}}(I - a\mathrm{v}\mathrm{v}^{\mathrm{T}}) \\ &= I - a\mathrm{v}\mathrm{v}^{\mathrm{T}} - a\mathrm{v}\mathrm{v}^{\mathrm{T}} + a^2\mathrm{v}\mathrm{v}^{\mathrm{T}}\mathrm{v}\mathrm{v}^{\mathrm{T}} \\ &= I - 2a\mathrm{v}\mathrm{v}^{\mathrm{T}} + a^2\mathrm{v}\mathrm{v}^{\mathrm{T}}\mathrm{v}\mathrm{v}^{\mathrm{T}}. \end{aligned} \tag{*}$$

Note that we have

$$\begin{aligned} \mathrm{v}\mathrm{v}^{\mathrm{T}}\mathrm{v}\mathrm{v}^{\mathrm{T}} &= \mathrm{v}(\mathrm{v}^{\mathrm{T}}\mathrm{v})\mathrm{v}^{\mathrm{T}} \\ &= \mathrm{v}\left(\frac{2}{a}\right)\mathrm{v}^{\mathrm{T}} && \text{by definition of } a \neq 0 \\ &= \frac{2}{a}\mathrm{v}\mathrm{v}^{\mathrm{T}}. \end{aligned}$$

Plugging this relation into (*), we obtain

$$AA = I - 2a\mathrm{v}\mathrm{v}^{\mathrm{T}} + a^2\frac{2}{a}\mathrm{v}\mathrm{v}^{\mathrm{T}} = I.$$

Thus we get $AA = I$.

This implies that the inverse matrix of $A$ is $A$ itself: $A^{-1} = A$.

**18** Consider the system of linear equations

$$\begin{aligned} x_1 &= 2, \\ -2x_1 + x_2 &= 3, \\ 5x_1 - 4x_2 + x_3 &= 2 \end{aligned}$$

**(a)** Find the coefficient matrix and its inverse matrix.
**(b)** Using the inverse matrix, solve the system of linear equations.
(*The Ohio State University*)

## Solution.

### (a) Find the coefficient matrix and its inverse matrix.

The coefficient matrix is

$$A := \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{bmatrix}.$$

To find the inverse, we reduce the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ 5 & -4 & 1 & 0 & 0 & 1 \end{array}\right]$$

Using the elementary row operation.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ 5 & -4 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3-5R_1]{R_2+2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & -4 & 1 & -5 & 0 & 1 \end{array}\right]$$
$$\xrightarrow{R_3+4R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 3 & 4 & 1 \end{array}\right].$$

Now that the left $3 \times 3$ part became the identity matrix (this form is called the reduced row echelon form), the inverse matrix is

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}.$$

### (b) Using the inverse matrix, solve the system of linear equations.

Using the coefficient matrix $A$ the given system can be written as the matrix equation

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}.$$

Multiplying it by the inverse matrix $A^{-1}$ on the left, we get

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= A^{-1} \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 20 \end{bmatrix}. \end{aligned}$$

Therefore the solution of the system is $x_1 = 2, x_2 = 7, x_3 = 20$.

## Comment.

Once you obtained the inverse matrix $A^{-1}$ in part (1), it's better to calculate $AA^{-1}$ using your matrix. If $AA^{-1}$ is not the identity matrix, you made a calculation mistake somewhere.

Also, in part (b), after finding the solution of the system, you may want to substitute your answer to the original system to see if they are actually the solution.

These extra (but not so time consuming) computations reduce careless mistakes in a exam.

⑲ Consider the following system of linear equations

$$\begin{aligned} 2x+3y+z&=-1\\ 3x+3y+z&=1\\ 2x+4y+z&=-2. \end{aligned}$$

**(a)** Find the coefficient matrix $A$ for this system.
**(b)** Find the inverse matrix of the coefficient matrix found in (a)
**(c)** Solve the system using the inverse matrix $A^{-1}$.

## Solution.

### (a) Find the coefficient matrix $A$ for this system.

The system can be written as

$$A\mathrm{x}=\mathrm{b},$$

where

$$A=\begin{bmatrix} 2 & 3 & 1\\ 3 & 3 & 1\\ 2 & 4 & 1 \end{bmatrix}$$

is the coefficient matrix of the system and

$$\mathrm{x}=\begin{bmatrix} x\\ y\\ z \end{bmatrix} \text{ and } \mathrm{b}=\begin{bmatrix} -1\\ 1\\ -2 \end{bmatrix}.$$

### (b) Find the inverse matrix of the coefficient matrix

We apply the elementary row operations to the augmented matrix $[A \mid I]$, where $I$ is the $3\times 3$ identity matrix.

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 1 & 0 \\ 2 & 4 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{R_2-R_1 \\ R_3-R_1}} \left[\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[\text{then } R_2 \leftrightarrow R_3]{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 2 & 3 & 1 & 1 & 0 & 0 \end{array}\right] \xrightarrow{R_3-2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 3 & 1 & 3 & -2 & 0 \end{array}\right]$$

$$\xrightarrow{R_3-3R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 6 & -2 & -3 \end{array}\right].$$

Now the left $3 \times 3$ part has become the identity matrix.

So the matrix $A$ is invertible and the inverse is given by the right $3 \times 3$ part.

Hence we obtain

$$A^{-1} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix}.$$

(c) Solve the system using the inverse matrix $A^{-1}$.

As noted in (a), the system can be written using matrices as

$$A\mathrm{x} = \mathrm{b}.$$

Multiplying by the inverse $A^{-1}$ on the left, we have

$$\begin{aligned} A^{-1}A\mathrm{x} &= A^{-1}\mathrm{b} \\ I\mathrm{x} &= A^{-1}\mathrm{b} \\ \mathrm{x} &= A^{-1}\mathrm{b}. \end{aligned}$$

Therefore the solution $\mathrm{x}$ of the system is given by

$$\begin{aligned} \mathrm{x} &= A^{-1}\mathrm{b} \\ &= \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}. \end{aligned}$$

Thus, the solution of the system is

$$x = 2, y = -1, z = -2.$$

**20** Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 4 \\ 1 & 1 & 0 \end{bmatrix}.$$

**(a)** Calculate the inverse matrix $A^{-1}$. If you think the matrix $A$ is not invertible, then explain why.
**(b)** Are the vectors

$$\mathbf{A}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}, \text{ and } \mathbf{A}_3 = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}$$

linearly independent?

**(c)** Write the vector $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as a linear combination of $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$.

(*The Ohio State University, Linear Algebra Exam*)

## Hint.

21 For (a), consider the augmented matrix $[A|I]$ and reduce it.

22 Note that given vectors are column vectors of the matrix $A$ .

23 Use the inverse matrix $A^{-1}$ to solve a system.

## Solution.

### (a) Calculate the inverse matrix $A^{-1}$

We consider the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 5 & 4 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

and reduce this matrix using the elementary row operations as follows.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 5 & 4 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3-R_1]{R_2-2R_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right] \xrightarrow[R_3+R_2]{R_1-2R_2}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 5 & -2 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -3 & 1 & 1 \end{array}\right] \xrightarrow[R_2-2R_3]{R_1+3R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 1 & 3 \\ 0 & 1 & 0 & 4 & -1 & -2 \\ 0 & 0 & 1 & -3 & 1 & 1 \end{array}\right].$$

Since the left $3 \times 3$ part of the last matrix is the identity matrix, the inverse matrix of $A$ is

$$A^{-1} = \begin{bmatrix} -4 & 1 & 3 \\ 4 & -1 & -2 \\ -3 & 1 & 1 \end{bmatrix}.$$

### (b) Are the Vectors Linearly Independent?

To check whether $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ are linearly independent, we consider the linear combination

$$x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + x_3\mathbf{A}_3 = 0$$

and if this equation has only zero solution, the vectors are linearly independent.

This equation can be written as the matrix equation

$$A\mathbf{x} = 0,$$

where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

Since by part (a), the inverse matrix $A^{-1}$ exists. Thus multiplying by $A^{-1}$ on the left we get $\mathbf{x} = \mathbf{0}$. Thus the solution is $\mathbf{x} = \mathbf{0}$ and the vectors $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ are linearly independent.

Another Solution of (b)

Note that $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ are columns vectors of the matrix $A$. We proved in part (a) that $A$ is invertible. We know that the column vectors of an invertible matrix are linearly independent. Thus the vectors $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ are linearly independent.

### (c) Write the vector $\mathbf{b}$ as a linear combination of $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$

We want to solve the vector equation

$$x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + x_3\mathbf{A}_3 = \mathbf{b}.$$

This can be written as the matrix equation

$$A\mathbf{x} = \mathbf{b}.$$

Since $A$ is invertible, we have

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -4 & 1 & 3 \\ 4 & -1 & -2 \\ -3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Thus $x_1 = 0, x_2 = 1$, and $x_3 = -1$ and the linear combination is

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{A}_2 - \mathbf{A}_3.$$

24 A $2 \times 2$ matrix $A$ satisfies $\operatorname{tr}(A^2) = 5$ and $\operatorname{tr}(A) = 3$. Find $\det(A)$.

### Solution 1 of (a) A $2 \times 2$ matrix $A$ satisfies $\operatorname{tr}(A^2) = 5$ and $\operatorname{tr}(A) = 3$. Find $\det(A)$.

Let

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then we have

$$A^2=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}=\begin{bmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{bmatrix}.$$

Since $\operatorname{tr}(A^2)=5$ and $\operatorname{tr}(A)=3$, we obtain

$$\begin{aligned} 5&=\operatorname{tr}(A^2)=(a^2+bc)+(bc+d^2)=a^2+2bc+d^2 \text{ and} \\ 3&=\operatorname{tr}(A)=a+d. \end{aligned}$$

We find the determinant $\det(A)=ad-bc$ as follows.

We have

$$\begin{aligned} \det(A)=ad-bc&=\frac{1}{2}\big((a+d)^2-(a^2+2bc+d^2)\big) \\ &=\frac{1}{2}(3^2-5)=2. \end{aligned}$$

Thus, we obtain $\det(A)=2$.

### Solution 2 of (a)

Let $\lambda_1$ and $\lambda_2$ be eigenvalues of $A$.

Then we have

$$\begin{aligned} 3=\operatorname{tr}(A)=\lambda_1+\lambda_2 \text{ and} \\ 5=\operatorname{tr}(A^2)=\lambda_1^2+\lambda_2^2. \end{aligned}$$

Here we used two facts.

The first one is that the trace of a matrix is the sum of all eigenvalues of the matrix.

The second one is that $\lambda^2$ is an eigenvalue of $A^2$ if $\lambda$ is an eigenvalue of $A$, and these are all the eigenvalues of $A^2$.

Since the determinant of $A$ is the product of eigenvalues of $A$, we have

$$\begin{aligned} \det(A)&=\lambda_1\lambda_2 \\ &=\frac{1}{2}\big((\lambda_1+\lambda_2)^2-(\lambda_1^2+\lambda_2^2)\big) \\ &=\frac{1}{2}(3^2-5) \\ &=2. \end{aligned}$$

Hence we have $\det(A)=2$.

### Solution 1 of (b) A $2\times 2$ matrix has two parallel columns and $\operatorname{tr}(A)=5$. Find $\operatorname{tr}(A^2)$.

Since two columns are parallel, we can write $A$ as

$$A = \begin{bmatrix} a & ra \\ c & rc \end{bmatrix}.$$

Then we have

$$5 = \operatorname{tr}(A) = a + rc.$$

We use the formula in Solution 1 of (a) for $\operatorname{tr}(A^2)$ with $b = ra$ and $d = rc$, and we compute

$$\begin{aligned} \operatorname{tr}(A^2) &= a^2 + 2(ra) + (rc)^2 \\ &= (a + rc)^2 \\ &= 5^2 = 25. \end{aligned}$$

Thus, we find $\operatorname{tr}(A^2) = 25$.

### Solution 2 of (b)

Since two columns are parallel, the matrix $A$ is singular. Hence $A$ has an eigenvalue $0$.

Since the sum of all the eigenvalues is $\operatorname{tr}(A) = 5$, we see that $0$ and $5$ are eigenvalues of $A$.

It follows that $0$ and $25$ are eigenvalues of $A^2$. Hence

$$\operatorname{tr}(A^2) = 0 + 25 = 25.$$

### Solution of (c) A $2 \times 2$ matrix $A$ has $\det(A) = 5$ and positive integer eigenvalues. What is the trace of $A$?

The product of eigenvalues of $A$ is the determinant $\det(A) = 5$.

Since eigenvalues are positive integers, it follows that $1$ and $5$ are eigenvalues of $A$.

It follows that

$$\operatorname{tr}(A) = 1 + 5 = 6.$$

(25) Suppose that a real matrix $A$ maps each of the following vectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

into the vectors

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{y}_2 = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{y}_3 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix},$$

respectively. That is, $A\mathbf{x}_i = \mathbf{y}_i$ for $i = 1, 2, 3$. Find the matrix $A$.
(*Kyoto University Exam*)

### Solution 1

By the condition that $A\mathbf{x}_i = \mathbf{y}_i$, we have a matrix equality $A[\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3] = [\mathbf{y}_1 \mathbf{y}_2 \mathbf{y}_3]$.

Explicitly we have

$$A\begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\end{bmatrix}=\begin{bmatrix}1 & -1 & 3\\2 & 0 & 1\\0 & 3 & 1\end{bmatrix}.$$

To find the matrix $A$, we find the inverse matrix of $\begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\end{bmatrix}$ and multiply on the right by it.

We use Gauss-Jordan elimination to transform the augmented matrix $[A|I]$ into $[I|A^{-1}]$.

We have

$$\left[\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & 0 & 0\\1 & 1 & 0 & 0 & 1 & 0\\1 & 1 & 1 & 0 & 0 & 1\end{array}\right]\xrightarrow[R_3-R_1]{R_2-R_1}\left[\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & 0 & 0\\0 & 1 & 0 & -1 & 1 & 0\\0 & 1 & 1 & -1 & 0 & 1\end{array}\right]$$
$$\xrightarrow{R_3-R_2}\left[\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & 0 & 0\\0 & 1 & 0 & -1 & 1 & 0\\0 & 0 & 1 & 0 & -1 & 1\end{array}\right].$$

Thus the inverse matrix is

$$\begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix}.$$

Hence we have

$$A=\begin{bmatrix}1 & -1 & 3\\2 & 0 & 1\\0 & 3 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix}=\begin{bmatrix}2 & -4 & 3\\2 & -1 & 1\\-3 & 2 & 1\end{bmatrix}.$$

## Solution 2

Note that in general $A$ is a matrix whose $i$-th column is the image of the $i$-th unit vector $\mathrm{e}_i$.

For given vectors $\mathrm{x}_i$, it is easy to see that

$$\begin{aligned}\mathrm{e}_1&=\mathrm{x}_1-\mathrm{x}_2\\\mathrm{e}_2&=\mathrm{x}_2-\mathrm{x}_3\\\mathrm{e}_3&=\mathrm{x}_3.\end{aligned}$$

Thus we have

$$A\mathrm{e}_1=A\mathrm{x}_1-A\mathrm{x}_2=\mathrm{y}_1-\mathrm{y}_2=\begin{bmatrix}2\\2\\-3\end{bmatrix}.$$

Similarly, we obtain

$$A\mathrm{e}_2=\begin{bmatrix}-4\\-1\\2\end{bmatrix},\text{ and }A\mathrm{e}_3=\begin{bmatrix}3\\1\\1\end{bmatrix}.$$

Therefore we have

$$A = [A\mathbf{e}_1\, A\mathbf{e}_2\, A\mathbf{e}_3] = \begin{bmatrix} 2 & -4 & 3 \\ 2 & -1 & 1 \\ -3 & 2 & 1 \end{bmatrix}.$$

Of course, this agrees with the answer of Solution 1.

## Comment.

This is a typical problem in linear algebra, but if you noticed the second approach then you could save time in the exam.

Once you obtained the matrix $A$, you can check whether you computed right by calculating $A\mathbf{x}_i$ with your $A$ and see if it is equal to the vector $\mathbf{y}_i$.

26 Let $A$ and $B$ are $n \times n$ matrices with real entries. Assume that $A+B$ is invertible. Then show that

$$A(A+B)^{-1}B = B(A+B)^{-1}A.$$

(*University of California, Berkeley Qualifying Exam*)

## Proof.

Let $P = A + B$. Then $B = P - A$.

Using these, we express the given expression in terms of only $A$ and $P$.
On one hand, we have

$$A(A+B)^{-1}B = AP^{-1}(P-A) = AP^{-1}P - AP^{-1}A = A - AP^{-1}A.$$

On the other hand we have

$$B(A+B)^{-1}A = (P-A)P^{-1}A = PP^{-1}A - AP^{-1}A = A - AP^{-1}A.$$

Thus these are equal.

This completes the proof.

27 Let $A$ be an $n \times n$ invertible matrix. Prove that the inverse matrix of $A$ is uniques.

## Hint.

That the inverse matrix of $A$ is unique means that there is only one inverse matrix of $A$.
(That's why we say "the" inverse matrix of $A$ and denote it by $A^{-1}$.)
So to prove the uniqueness, suppose that you have two inverse matrices $B$ and $C$ and show that in fact $B = C$.

Recall that $B$ is the inverse matrix if it satisfies

$$AB = BA = I,$$

where $I$ is the identity matrix.

### Proof.

Suppose that there are two inverse matrices $B$ and $C$ of the matrix $A$. Then they satisfy

$$AB = BA = I \tag{*}$$

and

$$AC = CA = I. \tag{**}$$

To show that the uniqueness of the inverse matrix, we show that $B = C$ as follows. Let $I$ be the $n \times n$ identity matrix.
We have

$$\begin{aligned} B &= BI \\ &= B(AC) && \text{by (**)} \\ &= (BA)C && \text{by the associativity} \\ &= IC && \text{by (*)} \\ &= C. \end{aligned}$$

Thus, we must have $B = C$, and there is only one inverse matrix of $A$.

㉘ Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$, and let $I$ be the $n \times n$ identity matrix. Suppose that the inner product of $\mathbf{u}$ and $\mathbf{v}$ satisfies

$$\mathbf{v}^{\mathrm{T}}\mathbf{u} \neq -1.$$

Define the matrix

$$A = I + \mathbf{u}\mathbf{v}^{\mathrm{T}}.$$

Prove that $A$ is invertible and the inverse matrix is given by the formula

$$A^{-1} = I - a\mathbf{u}\mathbf{v}^{\mathrm{T}},$$

where

$$a = \frac{1}{1 + \mathbf{v}^{\mathrm{T}}\mathbf{u}}.$$

This formula is called the **Sherman-Woodberry formula**.

### Proof.

Let us put

$$B = I - a\mathrm{uv}^{\mathrm{T}},$$

the matrix given by the Sherman-Woodberry formula.
We compute $AB$ and $BA$ and show that they are equal to the identity matrix $I$.

Let us first compute the matrix product $AB$. We have

$$\begin{aligned} AB &= (I+\mathrm{u}\mathrm{v}^{\mathrm{T}})(I-a\mathrm{u}\mathrm{v}^{\mathrm{T}}) \\ &= I - a\mathrm{u}\mathrm{v}^{\mathrm{T}} + \mathrm{u}\mathrm{v}^{\mathrm{T}} - a\mathrm{u}\mathrm{v}^{\mathrm{T}}\mathrm{u}\mathrm{v}^{\mathrm{T}} \\ &= I + (1-a)\mathrm{u}\mathrm{v}^{\mathrm{T}} - a\mathrm{u}\mathrm{v}^{\mathrm{T}}\mathrm{u}\mathrm{v}^{\mathrm{T}} \qquad (*) \end{aligned}$$

By using the defining formula for $a = \frac{1}{1+\mathrm{v}^{\mathrm{T}}\mathrm{u}}$, we have

$$a(1+\mathrm{v}^{\mathrm{T}}\mathrm{u}) = 1,$$

and thus

$$1 - a = a\mathrm{v}^{\mathrm{T}}\mathrm{u}. \qquad (**)$$

---

Note that the third term in (*) $-a\mathrm{u}\mathrm{v}^{\mathrm{T}}\mathrm{u}\mathrm{v}^{\mathrm{T}}$ contains $\mathrm{v}^{\mathrm{T}}\mathrm{u}$ in the middle, and $\mathrm{v}^{\mathrm{T}}\mathrm{u}$ is just a number. Thus we can factor out this number and get

$$\begin{aligned} -a\mathrm{u}\mathrm{v}^{\mathrm{T}}\mathrm{u}\mathrm{v}^{\mathrm{T}} &= -a\mathrm{u}(\mathrm{v}^{\mathrm{T}}\mathrm{u})\mathrm{v}^{\mathrm{T}} \\ &= -a(\mathrm{v}^{\mathrm{T}}\mathrm{u})\mathrm{u}\mathrm{v}^{\mathrm{T}} \qquad (***) \end{aligned}$$

Inserting (**) and (***) into (*), it follows that we have

$$\begin{aligned} AB &= I + (a\mathrm{v}^{\mathrm{T}}\mathrm{u})\mathrm{u}\mathrm{v}^{\mathrm{T}} - a(\mathrm{v}^{\mathrm{T}}\mathrm{u})\mathrm{u}\mathrm{v}^{\mathrm{T}} \\ &= I. \end{aligned}$$

Thus we have proved $AB = I$ .

---

Now we compute $BA$ . We have

$$\begin{aligned} BA &= (I-a\mathrm{u}\mathrm{v}^{\mathrm{T}})(I+\mathrm{u}\mathrm{v}^{\mathrm{T}}) \\ &= I + \mathrm{u}\mathrm{v}^{\mathrm{T}} - a\mathrm{u}\mathrm{v}^{\mathrm{T}} - a\mathrm{u}\mathrm{v}^{\mathrm{T}}\mathrm{u}\mathrm{v}^{\mathrm{T}} \\ &= I + (1-a)\mathrm{u}\mathrm{v}^{\mathrm{T}} - a\mathrm{u}\mathrm{v}^{\mathrm{T}}\mathrm{u}\mathrm{v}^{\mathrm{T}} \end{aligned}$$

and this is exactly the expression (*), hence $BA = AB = I$ .

---

Therefore, we conclude that the matrix $A$ is invertible and the inverse matrix is $B$. Hence

$$A^{-1} = I - a\mathrm{u}\mathrm{v}^{\mathrm{T}}$$

and we have proved the Sherman-Woodberry formula.

## Comment.

The invertible matrix theorem says that once we have $AB = I$ , then we have automatically $BA = I$ and the inverse matrix of $A$ is $B$, that is, $A^{-1} = B$.
So in the above proof, after proving $AB = I$ , you may conclude that $A$ is invertible and $A^{-1} = B$.

㉙ Let $A$ be a singular $2 \times 2$ matrix such that $\operatorname{tr}(A) \neq -1$ and let $I$ be the $2 \times 2$ identity matrix. Then prove that the inverse matrix of the matrix $I + A$ is given by the following formula:

$$(I+A)^{-1} = I - \frac{1}{1+\operatorname{tr}(A)}A.$$

Using the formula, calculate the inverse matrix of $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

## Proof.

We have

$$\begin{aligned}(I+A)\left(I - \frac{1}{1+\operatorname{tr}(A)}A\right) &= I - \frac{1}{1+\operatorname{tr}(A)}A + A - \frac{1}{1+\operatorname{tr}(A)}A^2 \\ &= I - \frac{1}{1+\operatorname{tr}(A)}\left(A - (1+\operatorname{tr}(A))A + A^2\right) \\ &= I - \frac{1}{1+\operatorname{tr}(A)}\left(A^2 - \operatorname{tr}(A)A\right). \qquad (*)\end{aligned}$$

The Cayley-Hamilton theorem for $2 \times 2$ matrices yields that

$$A^2 - \operatorname{tr}(A)A + \det(A)I = O.$$

Since $A$ is singular, we have $\det(A) = 0$.

Hence it follows that we have

$$A^2 - \operatorname{tr}(A)A = O,$$

and we obtain from (*) that

$$(I+A)\left(I - \frac{1}{1+\operatorname{tr}(A)}A\right) = I.$$

Similarly,

$$\left(I - \frac{1}{1+\operatorname{tr}(A)}A\right)(I+A) = I.$$

Therefore, we conclude that the inverse matrix of $I + A$ is given by the formula

$$(I+A)^{-1} = I - \frac{1}{1+\operatorname{tr}(A)}A.$$

### Find the inverse matrix of $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ using the formula

Now let us find the inverse matrix of $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ using the formula.

We first write

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = I + A,$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then $A$ is a singular matrix with $\operatorname{tr}(A) = 2$.

The formula yields that

$$\begin{aligned} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} &= (I + A)^{-1} \\ &= I - \frac{1}{3}A \\ &= \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}. \end{aligned}$$

## Related Question.

There is a similar formula for inverse matrices of certain $n \times n$ matrices, called **Sherman-Woodberry formula**.

See the post ↴

Sherman-Woodbery Formula for the Inverse Matrix

for the statement of the Sherman-Woodberry formula and its proof.

㉚ Let $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$. Show that

**(a)** $A^n = \begin{bmatrix} a^n & 0 \\ 0 & b^n \end{bmatrix}$ for any $n \in \mathbb{N}$.

**(b)** Let $B = S^{-1}AS$, where $S$ be an invertible $2 \times 2$ matrix. Show that $B^n = S^{-1}A^nS$ for any $n \in \mathbb{N}$

## Hint.

Use mathematical induction.

## Proof.

**(1)** We prove $A^n = \begin{bmatrix} a^n & 0 \\ 0 & b^n \end{bmatrix}$ by induction on $n$.

The base case $n = 1$ is true by definition.

Suppose that $A^k = \begin{bmatrix} a^k & 0 \\ 0 & b^k \end{bmatrix}$. Then we have

$$A^{k+1} = AA^k = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} a^k & 0 \\ 0 & b^k \end{bmatrix} = \begin{bmatrix} a^{k+1} & 0 \\ 0 & b^{k+1} \end{bmatrix}.$$

Here we used the induction hypothesis in the second equality.

Hence the inductive step holds. This completes the proof.

**(2)** We show that $B^n = S^{-1}A^nS$ by induction on $n$.
When $n = 1$, this is just the definition of $B$.

For induction step, assume that $B^k = S^{-1}A^kS$.
Then we have

$$B^{k+1} = BB^k = (S^{-1}AS)(S^{-1}A^kS) = S^{-1}AA^kS = S^{-1}A^{k+1}S,$$

where we used the induction hypothesis in the second equality and the third equality follows by canceling $SS^{-1} = I_2$ in the middle.

Thus the inductive step holds, and this competes the proof.

**31** Let $A$ be an $n \times n$ invertible matrix. Then prove the transpose $A^{\mathrm{T}}$ is also invertible and that the inverse matrix of the transpose $A^{\mathrm{T}}$ is the transpose of the inverse matrix $A^{-1}$. Namely, show that $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$.

## Proof.

By definition of inverse matrices, if there is an $n \times n$ matrix $B$ such that

$$A^{\mathrm{T}}B = I \text{ and } BA^{\mathrm{T}} = I,$$

where $I$ is the $n \times n$ identity matrix, then $A^{\mathrm{T}}$ is invertible and its inverse is $B$, that is, $B = (A^{\mathrm{T}})^{-1}$.

We claim that we can take $(A^{-1})^{\mathrm{T}}$ for this $B$.
In fact, we have

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = I^{\mathrm{T}} = I.$$

Here in the first equality, we used the fact about transpose matrices that

$$(CD)^{\mathrm{T}} = D^{\mathrm{T}}C^{\mathrm{T}}$$

for any matrices $C, D$ such that the matrix product $CD$ is well-defined.

Similarly, we have

$$(A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = I^{\mathrm{T}} = I.$$

This proves the transpose $A^{\mathrm{T}}$ is invertible and that $(A^{-1})^{\mathrm{T}}$ is the inverse matrix of $A^{\mathrm{T}}$, that is,

$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$$

as required.

**32** A square matrix $A$ is called ***nilpotent*** if there exists a positive integer $k$ such that $A^k = O$, where $O$ is the zero matrix.
**(a)** If $A$ is a nilpotent $n \times n$ matrix and $B$ is an $n \times n$ matrix such that $AB = BA$. Show that the product $AB$ is nilpotent.
**(b)** Let $P$ be an invertible $n \times n$ matrix and let $N$ be a nilpotent $n \times n$ matrix. Is the product $PN$ nilpotent? If so, prove it. If not, give a counterexample.

## Hint.

For (b), the statement is false. Try to find a counter example.

A typical nilpotent matrix is an upper triangular matrix whose diagonal entries are all zero.

## Proof.

### (a) Show that $AB$ is nilpotent

Since $A$ is nilpotent, there exists a positive integer $k$ such that $A^k = O$. Then we have

$$(AB)^k = (AB)(AB)\cdots(AB) = A^k B^k = OB^k = O.$$

Here in the second equality, we used the assumption that $AB = BA$ .

Thus we have $(AB)^k = O$ , hence the product matrix $AB$ is nilpotent.

### (b) Is $PN$ nilpotent?

In general, the product $PN$ of an invertible matrix $P$ and a nilpotent matrix $N$ is not nilpotent.

Here is a counterexample.

Let

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } N = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then the matrix $P$ is invertible since $\det(P) = 1$.

(Note that $P$ is a lower triangular matrix. So the determinant is the product of diagonal entries.)

---

Also, it is easy to see by direct computation that $N^3 = O$, hence $N$ is nilpotent. Indeed,

$$N^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$N^3 = N^2 N = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = O.$$

---

Now the product $PN$ is

$$PN = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

We show that $PN$ is not nilpotent.

We have

$$(PN)^2 = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$(PN)^3 = (PN)^2(PN) = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

This calculation shows that

$$(PN)^k = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \neq O \text{ for all } k \geq 2.$$

Thus $PN$ is not nilpotent. In conclusion, the product $PN$ of the invertible matrix $P$ and the nilpotent matrix $N$ is not nilpotent.

㉝ **(a)** Show that if $A$ is invertible, then $A$ is nonsingular.
**(b)** Let $A, B, C$ be $n \times n$ matrices such that $AB = C$. Prove that if either $A$ or $B$ is singular, then so is $C$.
**(c)** Show that if $A$ is nonsingular, then $A$ is invertible.

---

## Definition (Nonsingular Matrix)

An $n \times n$ matrix $A$ is called **nonsingular** if the equation $A\mathbf{x} = \mathbf{0}$ has only the zero solution $\mathbf{x} = \mathbf{0}$.

For basic properties of a nonsingular matrix, see the problem Properties of nonsingular and singular matrices.
The result of this problem will be used in the proof below.

## Proof.

### (a) If $A$ is invertible, then $A$ is nonsingular

Suppose that $A$ is invertible. This means that we have the inverse matrix $A^{-1}$ of $A$.
Consider the equation $A\mathbf{x} = \mathbf{0}$. We show that this equation has only zero solution.

Multiplying it by $A^{-1}$ on the left, we obtain

$$\begin{aligned} A^{-1}A\mathbf{x} &= A^{-1}\mathbf{0} \\ \Rightarrow \mathbf{x} &= \mathbf{0} \end{aligned}$$

Hence $A$ is nonsingular.

### (b) If either $A$ or $B$ is singular, then so is $C$

Suppose first that the matrix $B$ is singular.
Then there exists nonzero vector $\mathbf{b} \neq \mathbf{0}$ such that $B\mathbf{b} = \mathbf{0}$. Then we have $C\mathbf{b} = AB\mathbf{b} = A\mathbf{0} = \mathbf{0}$.

Since $\mathbf{b}$ is a nonzero vector, the matrix $C$ is singular.

Next, we assume that $B$ is nonsingular and $A$ is singular.
Since $A$ is singular, there exists nonzero vector $\mathbf{y}_0$ such that $A\mathbf{y}_0 = 0$.

Then consider the equation $B\mathbf{x} = \mathbf{y}_0$ has a unique solution $\mathbf{x}_0$ by part (c) of Properties of nonsingular and singular matrices.
The vector $\mathbf{x}_0$ is nonzero because $\mathbf{y}_0$ is nonzero.

Then we have

$$C\mathbf{x}_0 = AB\mathbf{x}_0 = A\mathbf{y}_0 = 0.$$

Therefore $C\mathbf{x} = 0$ has nonzero solution $\mathbf{x}_0$, hence $C$ is singular.

### (c) If $A$ is nonsingular, then $A$ is invertible

Suppose that $A$ is nonsingular. Let $\mathbf{e}_i$ be the $n$-dimensional vector whose entries are all $0$ but $1$ in the $i$th place.
Then the equations $A\mathbf{x} = \mathbf{e}_i$ has a unique solution $\mathbf{x}_i$ for $i = 1, \dots, n$.

Create a matrix $B$ whose $i$-th column vector is $\mathbf{x}_i$, namely $B = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_n]$.
Then we have $AB = [\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_n] = I_n$.

Hence $B$ is the right inverse of $A$.
Note that the identity matrix $I_n$ is nonsingular. Thus by part (c) of Properties of nonsingular and singular matrices, the matrix $B$ must be nonsingular as well.

We repeat the above argument using $B$ instead of $A$. Then there exist a matrix $C$ such that $BC = I_n$. We claim that $C = A$.
To see this, multiply $AB = I_n$ by $C$ on the right, we get $ABC = C$.

Since $BC = I_n$, we get $A = C$.

In summary, we obtained $AB = I_n$ and $BA = I_n$. Thus $A$ is invertible with the inverse $B$.

㉞ A square matrix $A$ is called **nilpotent** if some power of $A$ is the zero matrix. Namely, $A$ is nilpotent if there exists a positive integer $k$ such that $A^k = O$, where $O$ is the zero matrix. Suppose that $A$ is a nilpotent matrix and let $B$ be an invertible matrix of the same size as $A$. Is the matrix $B - A$ invertible? If so prove it. Otherwise, give a counterexample.

## Solution.

We claim that the matrix $B - A$ is not necessarily invertible.
Consider the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

This matrix is nilpotent as we have

$$A^2 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Also consider the matrix

$$B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Since the determinant of the matrix $B$ is $1$, it is invertible.

So the matrix $A$ and $B$ satisfy the assumption of the problem.

However the matrix

$$B - A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

is not invertible as its determinant is $0$.

Hence we found a counterexample.

## Related Question.

Here is another problem about a nilpotent matrix.

**Problem.**

Let $A$ be an $n \times n$ nilpotent matrix. Then prove that $I - A, I + A$ are both nonsingular matrices, where $I$ is the $n \times n$ identity matrix.

㉟ Let $A$ be an $n \times n$ matrix. The $(i, j)$ **cofactor** $C_{ij}$ of $A$ is defined to be $C_{ij} = (-1)^{ij} \det(M_{ij})$, where $M_{ij}$ is the $(i, j)$ **minor matrix** obtained from $A$ removing the $i$-th row and $j$-th column. Then consider the $n \times n$ matrix $C = (C_{ij})$, and define the $n \times n$ matrix $\operatorname{Adj}(A) = C^{\mathrm{T}}$. The matrix $\operatorname{Adj}(A)$ is called the **adjoint** matrix of $A$. When $A$ is invertible, then its inverse can be obtained by the formula

$$A^{-1} = \frac{1}{\det(A)} \operatorname{Adj}(A).$$

For each of the following matrices, determine whether it is invertible, and if so, then find the invertible matrix using the above formula.

**(a)** $A = \begin{bmatrix} 1 & 5 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$

**(b)** $B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 3 & 0 & 1 \end{bmatrix}.$

## Solution.

### (a) The Inverse Matrix of $A$.

Since $A$ is an upper triangular matrix, the determinant of $A$ is the product of diagonal entries.

Thus we have $\det(A) = -1 \neq 0$, and hence $A$ is invertible.

To find the inverse using the formula, we first determine the cofactors $C_{ij}$ of $A$.

We have

$$\begin{aligned}
C_{11} &= \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -1, \quad C_{12} = -\begin{vmatrix} 0 & 2 \\ 0 & 1 \end{vmatrix} = 0, \quad C_{13} = \begin{vmatrix} 0 & -1 \\ 0 & 0 \end{vmatrix} = 0 \\
C_{21} &= -\begin{vmatrix} 5 & 2 \\ 0 & 1 \end{vmatrix} = -5, \quad C_{22} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1, \quad C_{23} = -\begin{vmatrix} 1 & 5 \\ 0 & 0 \end{vmatrix} = 0 \\
C_{31} &= \begin{vmatrix} 5 & 2 \\ -1 & 2 \end{vmatrix} = 12, \quad C_{32} = -\begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = -2, \quad C_{33} = \begin{vmatrix} 1 & 5 \\ 0 & -1 \end{vmatrix} = -1.
\end{aligned}$$

The the adjoint matrix of $A$ is

$$\operatorname{Adj}(A) = C^{\mathrm{T}} = \begin{bmatrix} -1 & -5 & 12 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{bmatrix}.$$

Using the formula, we obtain the inverse matrix

$$A^{-1} = \frac{1}{\det(A)} \operatorname{Adj}(A) = \begin{bmatrix} 1 & 5 & -12 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

### (b) The Inverse Matrix of $B$.

To check the invertibility of the matrix $B$, we compute the determinant of $B$.

The second column cofactor expansion yields that

$$\det(B) = \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5 \neq 0.$$

So the matrix $B$ is invertible.

Now the cofactors $C_{ij}$ of $B$ are

$$\begin{aligned}
C_{11} &= \begin{vmatrix} 1 & 4 \\ 0 & 1 \end{vmatrix} = 1, \quad C_{12} = -\begin{vmatrix} 0 & 4 \\ 3 & 1 \end{vmatrix} = 12. \quad C_{13} = \begin{vmatrix} 0 & 1 \\ 3 & 0 \end{vmatrix} = -3 \\
C_{21} &= -\begin{vmatrix} 0 & 2 \\ 0 & 1 \end{vmatrix} = 0, \quad C_{22} = \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5, \quad C_{23} = -\begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} = 0 \\
C_{31} &= \begin{vmatrix} 0 & 2 \\ 1 & 4 \end{vmatrix} = -2, \quad C_{32} = -\begin{vmatrix} 1 & 2 \\ 0 & 4 \end{vmatrix} = -4, \quad C_{33} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.
\end{aligned}$$

Hence the adjoint matrix of $B$ is

$$\operatorname{Adj}(B) = C^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & -2 \\ 12 & -5 & -4 \\ -3 & 0 & 1 \end{bmatrix}.$$

It follows from the formula that the inverse matrix of $B$ is

$$B^{-1} = \frac{1}{\det(B)} \operatorname{Adj}(B) = \frac{1}{5}\begin{bmatrix} -1 & 0 & 2 \\ -12 & 5 & 4 \\ 3 & 0 & -1 \end{bmatrix}.$$

36 Let $A$ be a real symmetric matrix whose diagonal entries are all positive real numbers. Is it true that the all of the diagonal entries of the inverse matrix $A^{-1}$ are also positive? If so, prove it. Otherwise, give a counterexample.

### Solution.

The statement is in general false. We give a counterexample.

Let us consider the following $2 \times 2$ matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

The matrix $A$ satisfies the required conditions, that is, $A$ is symmetric and its diagonal entries are positive.

The determinant $\det(A) = (1)(1) - (2)(2) = -3$ and the inverse of $A$ is given by

$$A^{-1} = \frac{1}{-3}\begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix}$$

by the formula for the inverse matrix for $2 \times 2$ matrices.

This shows that the diagonal entries of the inverse matrix $A^{-1}$ are negative.

---

37 Let $A$ be an $n \times n$ nonsingular matrix with integer entries. Prove that the inverse matrix $A^{-1}$ contains only integer entries if and only if $\det(A) = \pm 1$.

### Hint.

38 If $B$ is a square matrix whose entries are integers, then the determinant of $B$ is an integer.

39 The inverse matrix of $A$ can be computed by the formula

$$A^{-1} = \frac{1}{\det(A)} \operatorname{Adj}(A).$$

### Proof.

Let $I$ be the $n \times n$ identity matrix.

**$(\implies)$: If $A^{-1}$ is an integer matrix, then $\det(A) = \pm 1$**

Suppose that every entry of the inverse matrix $A^{-1}$ is an integer.
It follows that $\det(A)$ and $\det(A^{-1})$ are both integers.

Since we have

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1,$$

we must have $\det(A) = \pm 1$.

> ($\Longleftarrow$): If $\det(A) = \pm 1$, then $A^{-1}$ is an integer matrix

Suppose that $\det(A) = \pm 1$. The inverse matrix of $A$ is given by the formula

$$A^{-1} = \frac{1}{\det(A)}\mathrm{Adj}(A),$$

where $\mathrm{Adj}(A)$ is the adjoint matrix of $A$.
Thus, we have

$$A^{-1} = \pm\mathrm{Adj}(A).$$

Note that each entry of $\mathrm{Adj}(A)$ is a cofactor of $A$, which is an integer.

(Recall that a cofactor is of the form $\pm\det(M_{ij})$, where $M_{ij}$ is the $(i,j)$ -minor matrix of $A$, hence entries of $M_{ij}$ are integers.)

Therefore, the inverse matrix $A^{-1}$ contains only integer entries.

# Subspaces in $\mathbb{R}^n$

**Definition**

Let $A$ be an $m \times n$ real matrix.

1. A subset $W$ in $\mathbb{R}^n$ is called a **subspace** if $W$ is a vector space in $\mathbb{R}^n$.
2. The **null space** $\mathcal{N}(A)$ of $A$ is defined by
$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}_m\}.$$
3. The **range** $\mathcal{R}(A)$ of the matrix $A$ is
$$\mathcal{R}(A) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$
4. The **column space** of $A$ is the subspace of $A^m$ spanned by the columns vectors of $A$.
5. The **row space** of $A$ is the subspace of $A^n$ spanned by the rows vectors of $A$.

**Summary**

Let $A$ be an $m \times n$ real matrix.

1. (**Subspace Criteria**) A subset $W$ in $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if and only if the following three condisions are met.
   **(a)** The zero vector $\mathbf{0} \in \mathbb{R}^n$ is in $W$.
   **(b)** If $\mathbf{x}, \mathbf{y} \in W$, then $\mathbf{x} + \mathbf{y} \in W$.
   **(c)** If $\mathbf{x} \in W$ and $c \in \mathbb{R}$, then $c\mathbf{x} \in W$.
2. The nullspace of $A$ is a subspace in $\mathbb{R}^n$.
3. The range of $A$ is a subspace in $\mathbb{R}^m$.
4. The range of $A$ is the column space of $A$.
5. If the columns of $A$ are linearly independent, then $\mathcal{N}(A) = \{\mathbf{0}\}$.

## Problems

1. Each of the following sets are not a subspace of the specified vector space. For each set, give a reason why it is not a subspace.

**(1)**

$$S_1 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 \geq 0 \right\}$$

in the vector space $\mathbb{R}^3$.

**(2)**

$$S_2 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 - 4x_2 + 5x_3 = 2 \right\}$$

in the vector space $\mathbb{R}^3$.

**(3)**

$$S_3 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; y = x^2 \right\}$$

in the vector space $\mathbb{R}^2$.

## Solution.

Recall the following subspace criteria.

A subset $W$ of a vector space $V$ over the scalar field $K$ is a subspace of $V$ if and only if the following three criteria are met.

2 The subset $W$ contains the zero vector of $V$.

3 If $u, v \in W$, then $u + v \in W$.

4 If $u \in W$ and $a \in K$, then $au \in W$.

Thus, to prove a subset $W$ is not a subspace, we just need to find a counterexample of any of the three criteria.

### Solution (1). $S_1 = \{\mathbf{x} \in \mathrm{R}^3 \mid x_1 \geq 0\}$

The subset $S_1$ does not satisfy condition 3. For example, consider the vector

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Then since $x_1 = 1 \geq 0$, the vector $\mathbf{x} \in S_1$. Then consider the scalar product of $\mathbf{x}$ and the scalar $-1$. Then we have

$$(-1) \cdot \mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix},$$

and the first entry is $-1$, hence $-\mathbf{x}$ is not in $S_1$. Thus $S_1$ does not satisfy condition 3 and it is not a subspace of $\mathrm{R}^3$.

(You can check that conditions 1, 2 are met.)

Solution (2). $S_2 = \{\mathbf{x} \in \mathrm{R}^3 \mid x_1 - 4x_2 + 5x_3 = 2\}$

The zero vector of the vector space $\mathrm{R}^3$ is

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Since the zero vector $\mathbf{0}$ does not satisfy the defining relation $x_1 - 4x_2 + 5x_3 = 2$, it is not in $S_2$. Hence condition 1 is not met, hence $S_2$ is not a subspace of $\mathrm{R}^3$.
(You can check that conditions 2, 3 are not met as well.)

Solution (3). $S_3 = \{\mathbf{x} \in \mathrm{R}^2 \mid y = x^2 \quad \}$

Consider vectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{and} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

These are vectors in $S_3$ since both vectors satisfy the defining relation $y = x^2$.

However, their sum

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

is not in $S_3$ since $1 \neq 0^2$.
Hence condition 2 is not met, and thus $S_3$ is not a subspace of $\mathrm{R}^2$.
(You can check that condition 1 is fulfilled yet condition 3 is not met.)

---

⑤ An $n \times n$ matrix $A$ is called **orthogonal** if $A^{\mathrm{T}}A = I$.
Let $V$ be the vector space of all real $2 \times 2$ matrices. Consider the subset

$$W := \{A \in V \mid A \text{ is an orthogonal matrix}\}.$$

Prove or disprove that $W$ is a subspace of $V$.

---

## Solution.

We claim that $W$ is not a subspace of $V$.

One way to see that $W$ is not a subspace of $V$ is to note that the zero vector $O$ in $V$, which is the $n \times n$ zero matrix, is not in $W$ as we have $O^{\mathrm{T}}O = O \neq I$.

Thus, $W$ is not a subspace of $V$.

### Another approach

You may also show that scalar multiplication (or addition) is not closed in $W$.

For example, the identity matrix $I$ is orthogonal as $I^{\mathrm{T}}I = I$, and thus $I$ is an element in $W$.
However, the scalar product $2I$ is not orthogonal since

$$(2I)^{\mathrm{T}}(2I) = 4I \neq I.$$

⑥ Let $A$ be an $m \times n$ matrix. Let $\mathcal{N}(A)$ be the null space of $A$. Suppose that $\mathbf{u} \in \mathcal{N}(A)$ and $\mathbf{v} \in \mathcal{N}(A)$. Let $\mathbf{w} = 3\mathbf{u} - 5\mathbf{v}$. Then find $A\mathbf{w}$.

### Hint.

Recall that the null space of an $m \times n$ matrix $A$ is a subspace of $\mathbb{R}^n$ defined by

$$\mathcal{N}(A) = \{\mathrm{x} \in \mathbb{R}^n \mid A\mathrm{x} = 0_m\}.$$

Here $0_m$ is the $m$-dimensional zero vector in $\mathbb{R}^m$.

### Solution.

Since $\mathrm{u}, \mathrm{v} \in \mathcal{N}(A)$, we have

$$A\mathrm{u} = 0_m \text{ and } A\mathrm{v} = 0_m,$$

where $0_m$ is the $m$-dimensional zero vector in $\mathbb{R}^m$.

Now using the properties of the matrix multiplication, we have

$$\begin{aligned} A\mathrm{w} &= A(3\mathrm{u} - 5\mathrm{v}) \\ &= A(3\mathrm{u}) + A(-5\mathrm{v}) \\ &= 3A\mathrm{u} - 5A\mathrm{v} \\ &= 30_m - 50_m = 0_m. \end{aligned}$$

Therefore we obtained

$$A\mathrm{w} = 0_m \in \mathbb{R}^m.$$

### Remark.

Note that a map $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathrm{x}) = A\mathrm{x}$, where $A$ is an $m \times n$ matrix is a linear transformation.
That is the map $T$ satisfies:

⑦ $T(\mathrm{u} + \mathrm{v}) = T(\mathrm{u}) + T(\mathrm{v})$ for any $\mathrm{u}, \mathrm{v} \in \mathbb{R}^n$, and

⑧ $T(c\mathrm{v}) = cT(\mathrm{v})$ for any $\mathrm{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

From this point of view, the above problem can be classified into a problem of linear transformation.

⑨ Let $W$ be the subset of the $3$-dimensional vector space $\mathbb{R}^3$ defined by

$$W=\left\{\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\in\mathbb{R}^3 \;\middle|\; 2x_1x_2=x_3\right\}.$$

**(a)** Which of the following vectors are in the subset $W$? Choose all vectors that belong to $W$.

$$(1)\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\quad (2)\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\quad (3)\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}\quad (4)\begin{bmatrix} 0 \\ 0 \end{bmatrix}\quad (5)\begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 & 4 \end{bmatrix}\quad (6)\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}.$$

**(b)** Determine whether $W$ is a subspace of $\mathbb{R}^3$ or not.

## Solution.

### Problem 1 (a) Which of the following vectors are in the subset $W$ ?

First of all, each element in $W$ is a $3$-dimensional vector. Thus (4), (5) can not be elements of $W$.

If $\mathbf{x}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, then $x_1=0, x_2=0, x_3=0$ and these satisfy the defining relation $2x_1x_2=x_3$. Thus $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is in $W$.

If $\mathbf{x}=\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, then we have $x_1=1, x_2=2, x_3=2$, and thus $2x_1x_2=4$ but $x_3=2$. Hence $2x_1x_2\neq x_3$, and $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ is not in $W$. For (3), we have $2x_1x_2=0=x_3$, and thus $\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$ is in $W$. For (6), we have $2x_1x_2=-2=x_3$, and thus $\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$ is in $W$.

In summary, the vectors (1), (3), (6) are in $W$.

### Problem 1 (b) Determine whether $W$ is a subspace of $\mathbb{R}^3$ or not.

We claim that $W$ is not a subspace. For example, consider the vectors

$$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}.$$

In part (a), we showed that these vectors are in $W$.

However, the sum

$$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}+\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}=\begin{bmatrix} 4 \\ -1 \\ -2 \end{bmatrix}$$

is not in $W$.

In fact, if $x_1=4, x_2=-1, x_3=-2$, then $2x_1x_2=-8\neq -2=x_3$.

Thus the sum does not satisfy the defining relation for $W$, hence not in $W$.

Therefore, the subset $W$ is not closed under addition, hence $W$ is not a subspace of $\mathbb{R}^3$.

## Solution of Problem 2

We prove that $W$ is a subspace of $\mathbb{R}^3$. To do this, we check the following subspace criteria.

**Subspace Criteria**

(a) The zero vector $\mathbf{0}\in\mathbb{R}^3$ is in $W$.

(b) If $\mathbf{x},\mathbf{y}\in W$, then $\mathbf{x}+\mathbf{y}\in W$.

(c) If $\mathbf{x}\in W$ and $c\in\mathbb{R}$, then $c\mathbf{x}\in W$.

---

Consider the zero vector $\mathbf{0}=\begin{bmatrix}0\\0\\0\end{bmatrix}\in\mathbb{R}^3$. Then $x_1=0, x_2=0, x_3=0$ satisfy the defining relations $x_1=3x_2$ and $x_3=0$. Thus the zero vector $\mathbf{0}$ is in $W$, hence condition (a) is satisfied.

---

To check condition (b), let us take $\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ and $\mathbf{y}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$ from $W$. Then we want to show that the sum $\mathbf{x}+\mathbf{y}$ is in $W$

Since $\mathbf{x}\in W$, we have

$$x_1=3x_2 \text{ and } x_3=0.$$

Also, since $\mathbf{y}\in W$, we have

$$y_1=3y_2 \text{ and } y_3=0.$$

From these, we see that

$$x_1+y_1=3x_2+3y_2=3(x_2+y_2) \text{ and } x_3+y_3=0.$$

Or simply we have

$$x_1+y_1=3(x_2+y_2) \text{ and } x_3+y_3=0.$$

Note that this implies that the sum

$$\mathbf{x}+\mathbf{y}=\begin{bmatrix}x_1+y_1\\x_2+y_2\\x_3+y_3\end{bmatrix}$$

satisfies the defining relation, and the sum is in $W$. Thus condition (b) is satisfied.

---

Finally, to prove condition (c), take $\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ from $W$ and let $c \in \mathbb{R}$. We need to show that the scalar product $c\mathbf{x} \in W$.

Since $\mathbf{x} \in W$, we have

$$x_1 = 3x_2 \text{ and } x_3 = 0.$$

Multiplying these by $c$, we have

$$cx_1 = 3(cx_2) \text{ and } cx_3 = 0.$$

These equalities tells that

the scalar product

$$c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ cx_3 \end{bmatrix}$$

satisfies the defining relation and hence it is in $W$. Thus condition (c) is met as well.

Since we have checked all subspace criteria, $W$ is a subspace of $\mathbb{R}^3$.

⑩ Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 4 \end{bmatrix}$ and let

$$\mathbf{a} = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

For each of the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, determine whether the vector is in the null space $\mathcal{N}(A)$. Do the same for the range $\mathcal{R}(A)$.

---

## Solution of (a)

Let us first recall the definitions of the null space and range. The null space of the $2 \times 3$ matrix $A$ is defined by

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^3 \mid A\mathbf{x} = 0\}$$

and the range of $A$ is defined by

$$\mathcal{R}(A) = \{\mathbf{y} \in \mathbb{R}^2 \mid A\mathbf{x} = \mathbf{y} \text{ for some } \mathbf{x} \in \mathbb{R}^3\}.$$

Note that the null space is a subspace of $\mathbb{R}^3$ and the range is a subspace of $\mathbb{R}^2$.

Consider first the vector $\mathbf{a} = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$. Since the vector $\mathbf{a}$ is three-dimensional, it cannot be in the range.

Also, since we have

$$A\mathbf{a} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 4 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq 0,$$

the vector $\mathbf{a}$ is not in the null space of $A$ .

---

Next, we consider the vector $\mathbf{b} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ . Again, this is three-dimensional vector, hence $\mathbf{b}$ is not in the range $\mathcal{R}(A)$ .

Since we have

$$A\mathbf{b} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0},$$

we see that the vector $\mathbf{b}$ is in the null space $\mathcal{N}(A)$ .

---

Finally, we consider the vector $\mathbf{c} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ . Since it is two-dimensional, the vector $\mathbf{c}$ is not in the null space $\mathcal{N}(A)$ .

Let us determine whether $\mathbf{c}$ is in the range $\mathcal{R}(A)$ . To do this, we check whether the system $A\mathbf{x} = \mathbf{c}$ is consistent or not. The augmented matrix of the system is

$$[A \mid \mathbf{c}] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 3 & 6 & 4 & 1 \end{array}\right].$$

Applying the elementary row operations, we have

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 3 & 6 & 4 & 1 \end{array}\right] \xrightarrow{R_2 - 3R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right] \xrightarrow{R_1 - R_2} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

and thus it has the general solution

$$\begin{aligned} x_1 &= -2x_2 + 3 \\ x_3 &= -2. \end{aligned}$$

Therefore, the vector $\mathbf{c}$ is in the range $\mathcal{R}(A)$ .

(For example, if $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$ , then $A\mathbf{x}_0 = \mathbf{c}$ .)

In summary, the vector $\mathbf{a}$ is neither in $\mathcal{N}(A)$ or in $\mathcal{R}(A)$ . The vector $\mathbf{b}$ is in $\mathcal{N}(A)$ but not in $\mathcal{R}(A)$ . The vector $\mathbf{c}$ is not in $\mathcal{N}(A)$ but in $\mathcal{R}(A)$ .

## Solution of (b)

We first describe the null space $\mathcal{N}(B)$ of the matrix $B$ .

The vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is in the null space $\mathcal{N}(B)$ if and only if $B\mathbf{x} = \mathbf{0}$ .

The augmented matrix of $B\mathbf{x} = \mathbf{0}$ is

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ -2 & -2 & -4 & 0 \end{array}\right].$$

Applying the elementary row operations, we have

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ -2 & -2 & -4 & 0 \end{array}\right] \xrightarrow{R_2+2R_1} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and the general solution is

$$x_1 = -x_2 - 2x_3.$$

Therefore, the vector $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is in the null space $\mathcal{N}(B)$ if and only if $x_1 = -x_2 - 2x_3$.

Thus,

$$\begin{aligned} \mathcal{N}(B) &= \{\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 = -x_2 - 2x_3\} \\ &= \{\mathrm{x} \in \mathbb{R}^3 \mid \mathrm{x} = \begin{bmatrix} -x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix}\} \\ &= \{\mathrm{x} \in \mathbb{R}^3 \mid \mathrm{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\} \\ &= \mathrm{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Thus, the set

$$S = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a spanning set of the null space $\mathcal{N}(B)$.

Let us verify that $S$ is a linearly independent set. Suppose that we have

$$c_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = 0,$$

for some scalars $c_1, c_2$. Then we have

$$\begin{bmatrix} -c_1 - 2c_2 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

and thus $c_1 = c_2 = 0$ and $S$ is linearly independent.

Hence $S$ is a linearly independent spanning set, and thus $S$ is a basis of the null space $\mathcal{N}(B)$.

⑪ Let $W$ be the subset of $\mathbb{R}^3$ defined by

$$W = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 = 3x_2 \text{ and } x_3 = 0 \right\}.$$

Determine whether the subset $W$ is a subspace of $\mathbb{R}^3$ or not.

## Solution.

### Problem 1 (a) Which of the following vectors are in the subset $W$ ?

First of all, each element in $W$ is a $3$-dimensional vector. Thus (4), (5) can not be elements of $W$.

If $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, then $x_1 = 0, x_2 = 0, x_3 = 0$ and these satisfy the defining relation $2x_1x_2 = x_3$. Thus $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is in $W$.

If $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, then we have $x_1 = 1, x_2 = 2, x_3 = 2$, and thus $2x_1x_2 = 4$ but $x_3 = 2$. Hence $2x_1x_2 \neq x_3$, and $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ is not in $W$. For (3), we have $2x_1x_2 = 0 = x_3$, and thus $\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$ is in $W$. For (6), we have $2x_1x_2 = -2 = x_3$, and thus $\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$ is in $W$.

In summary, the vectors (1), (3), (6) are in $W$.

### Problem 1 (b) Determine whether $W$ is a subspace of $\mathbb{R}^3$ or not.

We claim that $W$ is not a subspace. For example, consider the vectors

$$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}.$$

In part (a), we showed that these vectors are in $W$.

However, the sum

$$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ -2 \end{bmatrix}$$

is not in $W$.

In fact, if $x_1 = 4, x_2 = -1, x_3 = -2$, then $2x_1x_2 = -8 \neq -2 = x_3$.

Thus the sum does not satisfy the defining relation for $W$, hence not in $W$.

Therefore, the subset $W$ is not closed under addition, hence $W$ is not a subspace of $\mathbb{R}^3$.

### Solution of Problem 2

We prove that $W$ is a subspace of $\mathbb{R}^3$. To do this, we check the following subspace criteria.

**Subspace Criteria**
(a) The zero vector $\mathbf{0} \in \mathbb{R}^3$ is in $W$.
(b) If $\mathbf{x}, \mathbf{y} \in W$, then $\mathbf{x} + \mathbf{y} \in W$.
(c) If $\mathbf{x} \in W$ and $c \in \mathbb{R}$, then $c\mathbf{x} \in W$.

---

Consider the zero vector $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3$. Then $x_1 = 0, x_2 = 0, x_3 = 0$ satisfy the defining relations $x_1 = 3x_2$ and $x_3 = 0$. Thus the zero vector $\mathbf{0}$ is in $W$, hence condition (a) is satisfied.

---

To check condition (b), let us take $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ from $W$. Then we want to show that the sum $\mathbf{x} + \mathbf{y}$ is in $W$

Since $\mathbf{x} \in W$, we have

$$x_1 = 3x_2 \text{ and } x_3 = 0.$$

Also, since $\mathbf{y} \in W$, we have

$$y_1 = 3y_2 \text{ and } y_3 = 0.$$

From these, we see that

$$x_1 + y_1 = 3x_2 + 3y_2 = 3(x_2 + y_2) \text{ and } x_3 + y_3 = 0.$$

Or simply we have

$$x_1 + y_1 = 3(x_2 + y_2) \text{ and } x_3 + y_3 = 0.$$

Note that this implies that the sum

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}$$

satisfies the defining relation, and the sum is in $W$. Thus condition (b) is satisfied.

---

Finally, to prove condition (c), take $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ from $W$ and let $c \in \mathbb{R}$. We need to show that the scalar product $c\mathbf{x} \in W$.

Since $\mathbf{x} \in W$, we have

$$x_1 = 3x_2 \text{ and } x_3 = 0.$$

Multiplying these by $c$, we have

$$cx_1 = 3(cx_2) \text{ and } cx_3 = 0.$$

These equalities tells that

the scalar product

$$c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ cx_3 \end{bmatrix}$$

satisfies the defining relation and hence it is in $W$. Thus condition (c) is met as well.

Since we have checked all subspace criteria, $W$ is a subspace of $\mathbb{R}^3$.

⑫ Let

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 2 \\ -1 & -3 & -4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 2 \\ 5 & 3 & 3 \end{bmatrix}.$$

Determine the null spaces of matrices $A$ and $B$.

## Proof.

### The null space of the matrix $A$

We first determine the null space of the matrix $A$.

By definition, the null space is

$$\mathcal{N}(A) := \{\mathbf{x} \in \mathbb{R}^3 \mid A\mathbf{x} = \mathbf{0}\},$$

that is, the null space of $A$ consists of the solution $\mathbf{x}$ of the linear system $A\mathbf{x} = \mathbf{0}$.

To solve the system $A\mathbf{x} = \mathbf{0}$, we apply the Gauss-Jordan elimination. We reduce the augmented matrix $[A|\mathbf{0}]$ by elementary row operations as follows.

$$\left[\begin{array}{rrr|r} 1 & 2 & 2 & 0 \\ 2 & 3 & 2 & 0 \\ -1 & -3 & -4 & 0 \end{array}\right] \xrightarrow[R_3+R_1]{R_2-2R_1} \left[\begin{array}{rrr|r} 1 & 2 & 2 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right] \xrightarrow[R_3-5R_2]{R_1+2R_2}$$
$$\left[\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{-R_2} \left[\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The last matrix is in reduced row echelon form, and the solutions $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ of $A\mathbf{x} = \mathbf{0}$ satisfy

$$x_1 = 2x_3 \text{ and } x_2 = -2x_3$$

and $x_3$ is a free variable.

Thus we have

$$\mathbf{x} = \begin{bmatrix} 2x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

for any $x_3$ are solutions.

Therefore the null space of the matrix $A$ is

$$\begin{aligned} \mathcal{N}(A) &= \left\{ \mathbf{x} \in \mathbb{R}^3 \;\middle|\; \mathbf{x} = x_3 \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \text{ for any } x_3 \in \mathbb{R} \right\} \\ &= \operatorname{Span}\left( \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \right), \end{aligned}$$

the subspace of $\mathbb{R}^3$ spanned by the vector $\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$.

## The null space of the matrix $B$

The same procedure works for the matrix $B$. Thus we omit some detail below.

For the matrix $B$, the augmented matrix is reduced to

$$\left[\begin{array}{ccc|c} 1 & 2 & 2 & 0 \\ 2 & 3 & 2 & 0 \\ 5 & 3 & 3 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

(Check the computation by yourself.)

This implies that $\mathbf{x} = 0$ is the only solution of the system $B\mathbf{x} = 0$.

Therefore, the null space $\mathcal{N}(B)$ of the matrix $B$ consists of just the zero vector:

$$\mathcal{N}(B) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

## Comment.

The dimension of the null space $\mathcal{N}(A)$ is $1$, and the dimension of the null space $\mathcal{N}(B)$ is $0$.

In other words, the nullity of the matrix $A$ is $1$, and the nullity of the matrix $B$ is $0$.

(Recall that the nullity of a matrix is just the dimension of the null space of the matrix.)

13 Fix the row vector $\mathbf{b} = \begin{bmatrix} -1 & 3 & -1 \end{bmatrix}$, and let $\mathbb{R}^3$ be the vector space of $3 \times 1$ column vectors. Define

$$W = \{\mathbf{v} \in \mathbb{R}^3 \mid \mathbf{b}\mathbf{v} = 0\}.$$

Prove that $W$ is a vector subspace of $\mathbb{R}^3$, and find a basis for $W$.

## Proof.

We verify the subspace criteria: the zero vector $\mathbf{0}$ of $\mathbb{R}^3$ is in $W$, and $W$ is closed under addition and scalar multiplication.

---

First, the zero element in $\mathbb{R}^3$ is $\mathbf{0}$, the $3 \times 1$ column vector whose entries are all $0$. Then clearly $\mathrm{b}\mathbf{0} = 0$, and so $\mathbf{0} \in W$.

---

Next, suppose $\mathrm{v}, \mathrm{w} \in W$, and $c \in \mathbb{R}$. Then $\mathrm{bv} = \mathrm{bw} = 0$, and so

$$\mathrm{b}(\mathrm{v}+\mathrm{w}) = \mathrm{bv} + \mathrm{bw} = 0.$$

Thus, $\mathrm{v} + \mathrm{w} \in W$.

---

Because, again, $\mathrm{bv} = 0$, we have

$$\mathrm{b}(c\mathrm{v}) = c\mathrm{bv} = c0 = 0.$$

Thus $c\mathrm{v} \in W$. These three criteria show that $W$ is a vector subspace of $\mathbb{R}^3$.

## Comment.

We can generalize the problem with an arbitrary $1 \times 3$ row vector $\mathrm{b}$.

The proof is almost identical.
(Look at the proof. We didn't use components of the row vector $\mathrm{b} = \begin{bmatrix} -1 & 3 & -1 \end{bmatrix}$.)

---

Note that vectors $\mathrm{u}, \mathrm{v} \in \mathbb{R}^3$ is said to be perpendicular if

$$\mathrm{u} \cdot \mathrm{v} = \mathrm{u}^{\mathrm{T}}\mathrm{v} = 0.$$

Thus, the result of the problem says that for a fixed vector $\mathrm{u} \in \mathbb{R}^3$, the set of vectors $\mathrm{v}$ that are perpendicular to $\mathrm{u}$ is a subspace in $\mathbb{R}^3$.
(Note that we appy the problem to $\mathrm{b} = \mathrm{u}^{\mathrm{T}}$.)

⑭ Let $W$ be the subset of $\mathbb{R}^3$ defined by

$$W = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; 5x_1 - 2x_2 + x_3 = 0 \right\}.$$

Exhibit a $1 \times 3$ matrix $A$ such that $W = \mathcal{N}(A)$, the null space of $A$. Conclude that the subset $W$ is a subspace of $\mathbb{R}^3$.

## Solution.

Note that the defining equation $5x_1 - 2x_2 + x_3 = 0$ can be written as

$$\begin{bmatrix} 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0.$$

Hence if we put $A = \begin{bmatrix} 5 & -2 & 1 \end{bmatrix}$, then the defining equation becomes

$$A\mathbf{x} = 0.$$

Hence, we have

$$W = \{x \in \mathbb{R}^3 \mid A\mathbf{x} = 0\},$$

which is exactly the null space of the $1 \times 3$ matrix $A$.
Hence we have proved that $W = \mathcal{N}(A)$.

In general, the null space of an $m \times n$ matrix is a subspace of the vector space $\mathbb{R}^n$.
(See the post The null space (the kernel) of a matrix is a subspace of $\mathbb{R}^n$.)
Since we showed that $W$ is the null space of $1 \times 3$ matrix $A$, we conclude that $W$ is a subspace of $\mathbb{R}^3$.

15 Let $V$ be a subset of the vector space $\mathbb{R}^n$ consisting only of the zero vector of $\mathbb{R}^n$. Namely $V = \{\mathbf{0}\}$. Then prove that $V$ is a subspace of $\mathbb{R}^n$.

## Proof.

To prove that $V = \{0\}$ is a subspace of $\mathbb{R}^n$, we check the following subspace criteria.

**Subspace Criteria**
(a) The zero vector $0 \in \mathbb{R}^n$ is in $V$.
(b) If $\mathbf{x}, \mathbf{y} \in V$, then $\mathbf{x} + \mathbf{y} \in V$.
(c) If $\mathbf{x} \in V$ and $c \in \mathbb{R}$, then $c\mathbf{x} \in V$.

Condition (a) is clear since $V$ consists of the zero vector $0$.

To check condition (b), note that the only element in $V = \{0\}$ is $0$. Thus if $\mathbf{x}, \mathbf{y} \in V$, then both $\mathbf{x}, \mathbf{y}$ are $0$. Hence

$$\mathbf{x} + \mathbf{y} = 0 + 0 = 0 \in V$$

and condition (b) is met.

To confirm condition (c), let $\mathbf{x} \in V$ and $c \in \mathbb{R}$. Then $\mathbf{x} = 0$.
We have

$$c\mathbf{x} = c0 = 0 \in V$$

and condition (c) is satisfied.

Hence we have checked all the subspace criteria, and hence the subset $V = \{0\}$ consisting only of the zero vector is a subspace of $\mathbb{R}^n$.

## What’s the dimension of the zero vector space?

What’s the dimension of the subspace $V = \{0\}$ ?

The dimension of a subspace is the number of vectors in a basis. So let us first find a basis of $V$.

Note that a basis of $V$ consists of vectors in $V$ that are linearly independent spanning set. Since $0$ is the only vector in $V$, the set $S = \{0\}$ is the only possible set for a basis.

However, $S$ is not a linearly independent set since, for example, we have a nontrivial linear combination $1 \cdot 0 = 0$.

Therefore, the subspace $V = \{0\}$ does not have a basis.
Hence the dimension of $V$ is zero.

16 Prove that the null space $\mathcal{N}(A)$ is a subspace of the vector space $\mathbb{R}^n$.

---

## Proof.

### Subspace criteria

We check the following criteria for $\mathcal{N}(A)$ to be a subspace of $\mathbb{R}^n$.

(a) The zero vector $0_n \in \mathbb{R}^n$ is in $\mathcal{N}(A)$.
(b) If $\mathbf{x}, \mathbf{y} \in \mathcal{N}(A)$, then $\mathbf{x} + \mathbf{y} \in \mathcal{N}(A)$.
(c) If $\mathbf{x} \in \mathcal{N}(A)$ and $c \in \mathbb{R}$, then $c\mathbf{x} \in \mathcal{N}(A)$.

### Check (a)

Since we have $A0_n = 0_m$, the zero vector $0_n \in \mathbb{R}^n$ is in $\mathcal{N}(A)$. Thus (a) is met.

### Check (b)

For (b), take arbitrary vectors $\mathbf{x}, \mathbf{y} \in \mathcal{N}(A)$.
By the definition of $\mathcal{N}(A)$, we have

$$A\mathbf{x} = 0_m \text{ and } A\mathbf{y} = 0_m. \tag{*}$$

We want to show that $\mathbf{x} + \mathbf{y} \in \mathcal{N}(A)$. Thus we need to show $A(\mathbf{x} + \mathbf{y}) = 0$.

This equality can be proved as follows.
We have

$$\begin{aligned} A(\mathbf{x} + \mathbf{y}) &= A\mathbf{x} + A\mathbf{y} \\ &= 0 + 0 = 0. \end{aligned}$$

Here the second equality follows from (*).
Thus $\mathbf{x} + \mathbf{y} \in \mathcal{N}(A)$ and (b) is satisfied.

Check (c)

To check (c), take arbitrary $\mathrm{x} \in \mathcal{N}(A)$ and a scalar $c \in \mathbb{R}$.
Since $\mathrm{x} \in \mathcal{N}(A)$, we have

$$A\mathrm{x} = 0_m .$$

Multiplying by the scalar $c$, we have

$$cA\mathrm{x} = c0_m .$$

This is equivalent to

$$A(c\mathrm{x}) = 0_m .$$

This equality implies that $c\mathrm{x} \in \mathcal{N}(A)$, hence (c) is met.

Thus we checked the conditions (a), (b), and (c) and conclude that the null space $\mathcal{N}(A)$ is a subspace of $\mathbb{R}^n$.

17 Prove that every plane in the 3-dimensional space $\mathbb{R}^3$ that passes through the origin is a subspace of $\mathbb{R}^3$.

Proof.

Each plane $P$ in $\mathbb{R}^3$ through the origin is given by the equation

$$ax + by + cz = 0$$

for some real numbers $a, b, c$.

That is, the plane $P$ is a set of vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ satisfying the equation $ax + by + cz = 0$:

$$P = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; ax + by + cz = 0 \right\}.$$

---

Now the equation can be written as the matrix equation

$$A\mathbf{x} = \mathbf{0},$$

where $A$ is the $1 \times 3$ matrix $A$, $\mathbf{x} \in \mathbb{R}^3$, and $\mathbf{0}$ is the $1$-dimensional zero vector given by

$$A = \begin{bmatrix} a & b & c \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \mathbf{0} = [0].$$

---

Thus, the plane can be written as

$$P = \{\mathbf{x} \in \mathbb{R}^3 \mid A\mathbf{x} = \mathbf{0}\}$$

and this is the definition of the nullspace $\mathcal{N}(A)$ of $A$. That is

$$P=\mathcal{N}(A).$$

---

Therefore, the plane $P$ is the nullspace of the $1 \times 3$ matrix $A$.

Since the nullspace of a matrix is always a subspace, we conclude that the plane $P$ is a subspace of $\mathbb{R}^3$ .

Therefore, every plane in $\mathbb{R}^3$ through the origin is a subspace of $\mathbb{R}^3$.

---

⑱

Let $S$ be the subset of $\mathbb{R}^4$ consisting of vectors $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$ satisfying $2x+3y+5z+7w=0$. Then prove that the set $S$ is a subspace of $\mathbb{R}^4$.

(*The Ohio State University*)

**Proof.**

First, in set theoretical notation, the definition of $S$ can be written as

$$S=\left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; 2x+3y+5z+7w=0 \right\}.$$

Let $A=\begin{bmatrix} 2 & 3 & 5 & 7 \end{bmatrix}$ be the $1 \times 4$ matrix. Then the defining equation $2x+3y+5z+7w=0$ can be written as

$$A\mathbf{x}=0,$$

where

$$\mathbf{x}=\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}.$$

It follows that the set $S$ is the null space of $A$, that is, $S=\mathcal{N}(A)$.

Since every null space is a subspace, we see that $S$ is also a subspace of $\mathbb{R}^4$

---

⑲

**(a)** Let $S$ be the subset of $\mathbb{R}^4$ consisting of vectors $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$ satisfying

$$2x+4y+3z+7w+1=0.$$

Determine whether $S$ is a subspace of $\mathbb{R}^4$. If so prove it. If not, explain why it is not a subspace.

**(b)** Let $S$ be the subset of $\mathbb{R}^4$ consisting of vectors $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$ satisfying

$$2x + 4y + 3z + 7w = 0.$$

Determine whether $S$ is a subspace of $\mathbb{R}^4$. If so prove it. If not, explain why it is not a subspace.
(*The Ohio State University*)

## Solution.

### (a) $2x + 4y + 3z + 7w + 1 = 0$

We claim that $S$ is not a subspace of $\mathbb{R}^4$.

If $S$ is a subspace of $\mathbb{R}^4$, then the zero vector $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ in $\mathbb{R}^4$ must lie in $S$.

However, the zero vector $\mathbf{0}$ does not satisfy the equation

$$2x + 4y + 3z + 7w + 1 = 0.$$

So $\mathbf{0} \notin S$, and we conclude that $S$ is not subspace of $\mathbb{R}^4$.

### (b) $2x + 4y + 3z + 7w = 0$

In a set theoretical notation, we have

$$S = \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; 2x + 4y + 3z + 7w = 0 \right\}.$$

Let $A$ be the $1 \times 4$ matrix defined by

$$A = \begin{bmatrix} 2 & 4 & 3 & 7 \end{bmatrix}.$$

Then the equation $2x + 4y + 3z + 7w = 0$ can be written as

$$A \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = 0.$$

So we have

$$\begin{aligned} S &= \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; A \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = 0 \right\} \\ &= \mathcal{N}(A), \end{aligned}$$

the null space of $A$ .

Recall that the null space of a matrix is always a subspace.
Hence the subset $S$ is a subspace of $\mathbb{R}^4$ as it is the null space of the matrix $A$ .

⑳ Let $\mathbb{R}^2$ be the $x$-$y$-plane. Then $\mathbb{R}^2$ is a vector space. A line $\ell \subset \mathbb{R}^2$ with slope $m$ and $y$-intercept $b$ is defined by

$$\ell = \{(x, y) \in \mathbb{R}^2 \mid y = mx + b\}.$$

Prove that $\ell$ is a subspace of $\mathbb{R}^2$ if and only if $b = 0$.

## Proof.

We must prove two statements. First we show that if $b \neq 0$, then $\ell$ is not a subspace. Then we show that if $b = 0$, then $\ell$ is a subspace.

In order for $\ell$ to be a subspace, it must contain the zero element $(0, 0)$. Plugging this point into the defining equation yields $b = 0$. Thus if $b \neq 0$, then $\ell$ cannot be a subspace.

Now we prove that if $b = 0$, then $\ell$ is a subspace. We have already shown that $(0, 0) \in \ell$. Now suppose we have two points $(x_1, y_1), (x_2, y_2) \in \ell$. Then we have

$$y_1 + y_2 = mx_1 + mx_2 = m(x_1 + x_2)$$

and so $(x_1 + x_2, y_1 + y_2)$ is contained in $\ell$.

Finally for $c \in R$ we must show that $c(x_1, y_1) = (cx_1, cy_1)$ lies in $\ell$. We check

$$cy_1 = c(mx_1) = m(cx_1),$$

and so $(cx_1, cy_1) \in \ell$. This proves that if $b = 0$, then $\ell$ is a subspace.

㉑ Let $S$ be the following subset of the 3-dimensional vector space $\mathbb{R}^3$.

$$S = \left\{ \mathbf{x} \in \mathbb{R}^3 \;\middle|\; \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, x_1, x_2, x_3 \in \mathbb{Z} \right\},$$

where $\mathbb{Z}$ is the set of all integers. Determine whether $S$ is a subspace of $\mathbb{R}^3$.

## Proof.

We claim that $S$ is not a subspace of $\mathbb{R}^3$.
If $S$ is a subspace, then $S$ is closed under scalar multiplication.
But this is not the case for $S$.

For example, consider $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Since all entries are integers, this is an element of $S$.

Let us compute the scalar multiplication of this vector $\mathbf{x}$ and the scalar $1/2 \in \mathbb{R}$. We have

$$\frac{1}{2} \cdot \mathbf{x} = \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}.$$

Since $1/2$ is not an integer, the scalar multiplication $\frac{1}{2}\mathbf{x}$ is not in $S$. Therefore the subset $S$ cannot be a subspace of $R^3$.

---

㉒ Let $\mathbf{a}$ and $\mathbf{b}$ be fixed vectors in $\mathbb{R}^3$, and let $W$ be the subset of $\mathbb{R}^3$ defined by

$$W = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{a}^{\mathrm{T}}\mathbf{x} = 0 \text{ and } \mathbf{b}^{\mathrm{T}}\mathbf{x} = 0\}.$$

Prove that the subset $W$ is a subspace of $\mathbb{R}^3$.

## Proof.

We prove the following criteria for the subset $W$ to be a subspace of $\mathbb{R}^3$.

**(a)** The zero vector $\mathbf{0} \in \mathbb{R}^3$ is in $W$.

**(b)** If $\mathbf{x}, \mathbf{y} \in W$, then $\mathbf{x} + \mathbf{y} \in W$.

**(c)** If $\mathbf{x} \in W$ and $c \in \mathbb{R}$, then $c\mathbf{x} \in W$.

For (a), note that $\mathbf{a}^{\mathrm{T}}\mathbf{0} = 0$ and $\mathbf{b}^{\mathrm{T}}\mathbf{0} = 0$. Thus the zero vector $\mathbf{0} \in \mathbb{R}^3$ is in $W$.

To check (b), let $\mathbf{x}, \mathbf{y} \in W$. Then we have the following relations.

$$\mathbf{a}^{\mathrm{T}}\mathbf{x} = 0 \text{ and } \mathbf{b}^{\mathrm{T}}\mathbf{x} = 0, \text{ and } \mathbf{a}^{\mathrm{T}}\mathbf{y} = 0 \text{ and } \mathbf{b}^{\mathrm{T}}\mathbf{y} = 0. \tag{*}$$

To show that $\mathbf{x} + \mathbf{y} \in W$, we need to show that

$$\mathbf{a}^{\mathrm{T}}(\mathbf{x} + \mathbf{y}) = 0 \text{ and } \mathbf{b}^{\mathrm{T}}(\mathbf{x} + \mathbf{y}) = 0.$$

We first compute

$$\begin{aligned} \mathbf{a}^{\mathrm{T}}(\mathbf{x} + \mathbf{y}) &= \mathbf{a}^{\mathrm{T}}\mathbf{x} + \mathbf{a}^{\mathrm{T}}\mathbf{y} \\ &= 0 + 0 = 0 \end{aligned}$$

by the relations (*).

Similarly, we have

$$\begin{aligned}\mathbf{b}^{\mathrm{T}}(\mathbf{x}+\mathbf{y}) &= \mathbf{b}^{\mathrm{T}}\mathbf{x}+\mathbf{b}^{\mathrm{T}}\mathbf{y}\\ &= 0+0=0\end{aligned}$$

by the relations (*).

Thus the vector $\mathbf{x}+\mathbf{y}$ satisfies the defining relations for $W$, hence $\mathbf{x}+\mathbf{y}\in W$.

---

Finally, to prove (c), let $\mathbf{x}\in W$ and let $c\in\mathbb{R}$.
Since $\mathbf{x}\in W$, we have

$$\mathbf{a}^{\mathrm{T}}\mathbf{x}=0 \text{ and } \mathbf{b}^{\mathrm{T}}\mathbf{x}=0.$$

Multiplying by the scalar $c$ from the left, we obtain

$$\mathbf{a}^{\mathrm{T}}(c\mathbf{x})=0 \text{ and } \mathbf{b}^{\mathrm{T}}(c\mathbf{x})=0.$$

(Note that since $c$ is a scalar, we can switch the order of the product of $c$ and $\mathbf{a}^{\mathrm{T}}$. Same for $c$ and $\mathbf{b}^{\mathrm{T}}$.)

These equalities proves that the vector $c\mathbf{x}$ satisfies the defining relation for $W$. Thus $c\mathbf{x}\in W$.

---

Therefore the criteria (a)-(c) are all met, and we conclude that $W$ is a subspace of $\mathbb{R}^3$.

---

(23) Let $U$ and $V$ be subspaces of the $n$-dimensional vector space $\mathbb{R}^n$. Prove that the intersection $U\cap V$ is also a subspace of $\mathbb{R}^n$.

---

## Definition (Intersection).

Recall that the intersection $U\cap V$ is the set of elements that are both elements of $U$ and $V$.
In the set theoretical notation, we have

$$U\cap V=\{x \mid x\in U \text{ and } x\in V\}.$$

## Proof.

To prove that the intersection $U\cap V$ is a subspace of $\mathbb{R}^n$, we check the following subspace criteria:

(24) The zero vector $\mathbf{0}$ of $\mathbb{R}^n$ is in $U\cap V$.

(25) For all $\mathbf{x},\mathbf{y}\in U\cap V$, the sum $\mathbf{x}+\mathbf{y}\in U\cap V$.

(26) For all $\mathbf{x}\in U\cap V$ and $r\in\mathbb{R}$, we have $r\mathbf{x}\in U\cap V$.

As $U$ and $V$ are subspaces of $\mathbb{R}^n$, the zero vector $0$ is in both $U$ and $V$.

Hence the zero vector $0 \in \mathbb{R}^n$ lies in the intersection $U \cap V$.

So condition 1 is met.

---

Suppose that $\mathbf{x}, \mathbf{y} \in U \cap V$.

This implies that $\mathbf{x}$ is a vector in $U$ as well as a vector in $V$.

Similarly, $\mathbf{y}$ is a vector in $U$ as well as a vector in $V$.

Since $U$ is a subspace and $\mathbf{x}$ and $\mathbf{y}$ are both vectors in $U$, their sum $\mathbf{x}+\mathbf{y}$ is in $U$.

Similarly, since $V$ is a subspace and $\mathbf{x}$ and $\mathbf{y}$ are both vectors in $V$, their sum $\mathbf{x}+\mathbf{y} \in V$.

Therefore the sum $\mathbf{x}+\mathbf{y}$ is a vector in both $U$ and $V$.

Hence $\mathbf{x}+\mathbf{y} \in U \cap V$.

Thus condition 2 is met.

---

To verify condition 3, let $\mathbf{x} \in U \cap V$ and $r \in \mathbb{R}$.

As $\mathbf{x} \in U \cap V$, the vector $\mathbf{x}$ lies in both $U$ and $V$.

Since both $U$ and $V$ are subspaces, the scalar multiplication is closed in $U$ and $V$, respectively.

Thus $r\mathbf{x} \in U$ and $r\mathbf{x} \in V$.

It follows that $r\mathbf{x} \in U \cap V$.

This proves condition 3, and hence the intersection $U \cap V$ is a subspace of $\mathbb{R}^n$.

27 Determine whether the following is true or false. If it is true, then give a proof. If it is false, then give a counterexample. Let $W_1$ and $W_2$ be subspaces of the vector space $\mathbb{R}^n$.If $B_1$ and $B_2$ are bases for $W_1$ and $W_2$, respectively, then $B_1 \cap B_2$ is a basis of the subspace $W_1 \cap W_2$.

## Solution.

The statement is false. We give a counterexample.

Let us consider the vector space $\mathbb{R}^2$, the plane.

Let $W_1$ and $W_2$ be $\mathbb{R}^2$ itself. Then they are subspaces of $\mathbb{R}^2$.

Let

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ and } B_2 = \left\{ \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$$

be bases of $W_1$ and $W_2$, respectively.

(Note. $B_1$ consists of two linearly independent vectors in the 2-dimensional vector space $W_1 = \mathbb{R}^2$, hence $B_1$ is a basis of $W_1$. Similarly, $B_2$ is a basis of $W_2$.)

Since $W_1$ and $W_2$ are both $\mathbb{R}^2$, we have $W_1 \cap W_2 = \mathbb{R}^2$.

However, the intersection $B_1 \cap B_2$ is the empty set, and the empty set is not a basis of $W_1 \cap W_2 = \mathbb{R}^2$. Thus, we have found a counterexample to the statement.

(28) Let $U$ and $V$ be subspaces of the vector space $\mathbb{R}^n$. If neither $U$ nor $V$ is a subset of the other, then prove that the union $U \cup V$ is not a subspace of $\mathbb{R}^n$.

## Proof.

Since $U$ is not contained in $V$, there exists a vector $\mathbf{u} \in U$ but $\mathbf{u} \notin V$.

Similarly, since $V$ is not contained in $U$, there exists a vector $\mathbf{v} \in V$ but $\mathbf{v} \notin U$.

Seeking a contradiction, let us assume that the union is $U \cup V$ is a subspace of $\mathbb{R}^n$.

The vectors $\mathbf{u}, \mathbf{v}$ lie in the vector space $U \cup V$.

Thus their sum $\mathbf{u} + \mathbf{v}$ is also in $U \cup V$.

This implies that we have either

$$\mathbf{u} + \mathbf{v} \in U \text{ or } \mathbf{u} + \mathbf{v} \in V.$$

If $\mathbf{u} + \mathbf{v} \in U$, then there exists $\mathbf{u}' \in U$ such that

$$\mathbf{u} + \mathbf{v} = \mathbf{u}'.$$

Since the vectors $\mathbf{u}$ and $\mathbf{u}'$ are both in the subspace $U$, their difference $\mathbf{u}' - \mathbf{u}$ is also in $U$. Hence we have

$$\mathbf{v} = \mathbf{u}' - \mathbf{u} \in U.$$

However, this contradicts the choice of the vector $\mathbf{v} \notin U$.

Thus, we must have $\mathbf{u} + \mathbf{v} \in V$.

In this case, there exists $\mathbf{v}' \in V$ such that

$$\mathbf{u} + \mathbf{v} = \mathbf{v}'.$$

Since both $\mathbf{v}, \mathbf{v}'$ are vectors of $V$, it follows that

$$\mathbf{u} = \mathbf{v}' - \mathbf{v} \in V,$$

which contradicts the choice of $\mathbf{u} \notin V$.

Therefore, we have reached a contradiction. Thus, the union $U \cup V$ cannot be a subspace of $\mathbb{R}^n$.

## Related Question.

In fact, the converse of this problem is true.

**Problem**. Let $W_1, W_2$ be subspaces of a vector space $V$. Then prove that $W_1 \cup W_2$ is a subspace of $V$ if and only if $W_1 \subset W_2$ or $W_2 \subset W_1$.

For a proof, see the post “Union of Subspaces is a Subspace if and only if One is Included in Another“.

㉙ Let $W_1, W_2$ be subspaces of a vector space $V$. Then prove that $W_1 \cup W_2$ is a subspace of $V$ if and only if $W_1 \subset W_2$ or $W_2 \subset W_1$.

## Proof.

### If $W_1 \cup W_2$ is a subspace, then $W_1 \subset W_2$ or $W_2 \subset W_1$.

( $\Longrightarrow$ ) Suppose that the union $W_1 \cup W_2$ is a subspace of $V$.
Seeking a contradiction, assume that $W_1 \not\subset W_2$ and $W_2 \not\subset W_1$.
This means that there are elements

$$x \in W_1 \setminus W_2 \text{ and } y \in W_2 \setminus W_1.$$

Since $W_1 \cup W_2$ is a subspace, it is closed under addition. Thus, we have $x+y \in W_1 \cup W_2$.

It follows that we have either

$$x+y \in W_1 \text{ or } x+y \in W_2.$$

Suppose that $x+y \in W_1$. Then we write

$$y=(x+y)-x.$$

Since both $x+y$ and $x$ are elements of the subspace $W_1$, their difference $y=(x+y)-x$ is also in $W_1$. However, this contradicts the choice of $y \in W_2 \setminus W_1$.

Similarly, when $x+y \in W_2$, then we have

$$x=(x+y)-y \in W_2,$$

and this contradicts the choice of $x \in W_1 \setminus W_2$.

In either case, we have reached a contradiction.
Therefore, we have either $W_1 \subset W_2$ or $W_2 \subset W_1$.

### If $W_1 \subset W_2$ or $W_2 \subset W_1$, then $W_1 \cup W_2$ is a subspace.

( $\Longleftarrow$ ) If we have $W_1 \subset W_2$, then it yields that $W_1 \cup W_2 = W_2$ and it is a subspace of $V$.

Similarly, if $W_2 \subset W_1$, then we have $W_1 \cup W_2 = W_2$ and it is a subspace of $V$.
In either case, the union $W_1 \cup W_2$ is a subspace.

㉚ Let $A$ and $B$ be $n \times n$ matrices. Then prove that

$$\mathcal{N}(A) \cap \mathcal{N}(B) \subset \mathcal{N}(A+B).$$

## Definition.

Recall that the **null space** (or **kernel**) of an $n \times n$ matrix is

$$\mathcal{N}(A) = \{\mathrm{x} \in \mathbb{R}^n \mid A\mathrm{x} = 0\}.$$

The null space $(\mathcal{N})$ is a subspace of the $n$-dimensional vector space $\mathbb{R}^n$.

## Proof.

Let $\mathbf{x}$ be an arbitrary vector in the intersection $\mathcal{N}(A) \cap \mathcal{N}(B)$.
Then the vector $\mathbf{x}$ belongs to both $\mathcal{N}(A)$ and $\mathcal{N}(B)$.
Thus, by definition of the null space, we have

$$A\mathrm{x} = 0 \text{ and } B\mathrm{x} = 0.$$

If follows from these equalities that we have

$$(A+B)\mathrm{x} = A\mathrm{x} + B\mathrm{x} = 0 + 0 = 0.$$

Hence $\mathbf{x}$ lies in the null space of the matrix $A + B$, that is, $\mathrm{x} \in \mathcal{N}(A+B)$.

Since $\mathbf{x}$ is an arbitrary element of $\mathcal{N}(A) \cap \mathcal{N}(B)$, we have shown the inclusion

$$\mathcal{N}(A) \cap \mathcal{N}(B) \subset \mathcal{N}(A+B),$$

as required.

# Bases and Dimension of Subspaces in $\mathbb{R}^n$

**Definition**

Let $V$ be a subspace in $\mathbb{R}^n$.

1. A basis for $V$ is a linearly independent spanning set for $V$.
2. The number of vectors in a basis for $V$ is called the **dimension** of $V$.
3. The dimension of the range $\mathcal{R}(A)$ of a matrix $A$ is called the **rank** of $A$.
4. The dimension of the null space $\mathcal{N}(A)$ of a matrix $A$ is called the **nullity** of $A$.

**Summary**

1. A basis is not unique.
2. The rank-nullity theorem:
   (Rank of $A$)+(Nullity of $A$)=(The number of columns in $A$).
3. The rank of $A$ and the rank of $A^{\mathrm{T}}$ are the same.

## Problems

1. Let

$$\mathbf{v}=\begin{bmatrix} a \\ b \\ c \end{bmatrix}, \qquad \mathbf{v}_1=\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2=\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}.$$

Find the necessary and sufficient condition so that the vector $\mathbf{v}$ is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2$.

### Solution 1. (Use the range)

The question is equivalent to finding the condition so that the vector $\mathbf{v}$ is in the range of the matrix

$$A=[\mathbf{v}_1, \mathbf{v}_2]=\begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 0 & 2 \end{bmatrix}.$$

The vector $\mathbf{v}$ is in the range $\mathcal{R}(A)$ if and only if the system $A\mathbf{x}=\mathbf{v}$ is consistent.

We reduce the augmented matrix of the system by elementary row operations as follows.

$$[A \mid \mathbf{v}] = \begin{bmatrix} 1 & 2 & a \\ 2 & -1 & b \\ 0 & 2 & c \end{bmatrix} \xrightarrow{R_2-2R_1} \begin{bmatrix} 1 & 2 & a \\ 0 & -5 & b-2a \\ 0 & 2 & c \end{bmatrix}$$

$$\xrightarrow{R_2+3R_3} \begin{bmatrix} 1 & 2 & a \\ 0 & 1 & b-2a+3c \\ 0 & 2 & c \end{bmatrix}$$

$$\xrightarrow{R_3-2R_2} \begin{bmatrix} 1 & 2 & a \\ 0 & 1 & b-2a+3c \\ 0 & 0 & 4a-2b-5c \end{bmatrix}.$$

The last matrix is in echelon form and the system is consistent if and only if $4a-2b-5c=0$.

Therefore, the condition that $\mathbf{v}$ be a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ is $4a-2b-5c=0$.

## Solution 2. (Use the cross product)

Note that the vectors $\mathbf{v}_1, \mathbf{v}_2$ spans a plane $P$ in $\mathrm{R}^3$.

Thus, the vector $\mathbf{v}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ if and only if $\mathbf{v}$ lies on the plane $P$.

The cross product

$$\mathbf{v}_1 \times \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \times \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} 2 & -1 \\ 0 & 2 \end{vmatrix} \\ \\ -\begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} \\ \\ \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ -5 \end{bmatrix}$$

is perpendicular to the plane $P$.

Therefore, the vector $\mathbf{v}$ is on the plane if and only if the dot (inner) product

$$\mathbf{v} \cdot (\mathbf{v}_1 \times \mathbf{v}_2) = 0.$$

Namely,

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -2 \\ -5 \end{bmatrix} = 4a-2b-5c=0,$$

and we obtained the same condition as in Solution 1.

---

② Find a basis of the null space of the matrix $B = \begin{bmatrix} 1 & 1 & 2 \\ -2 & -2 & -4 \end{bmatrix}$.

## Solution of (a)

Let us first recall the definitions of the null space and range. The null space of the 2 × 3 matrix $A$ is defined by

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathrm{R}^3 \mid A\mathbf{x} = \mathbf{0}\}$$

and the range of $A$ is defined by

$$\mathcal{R}(A) = \{\mathbf{y} \in \mathrm{R}^2 \mid A\mathbf{x} = \mathbf{y} \text{ for some } \mathbf{x} \in \mathrm{R}^3\}.$$

Note that the null space is a subspace of $\mathrm{R}^3$ and the range is a subspace of $\mathrm{R}^2$.

Consider first the vector $\mathbf{a} = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$. Since the vector $\mathbf{a}$ is three-dimensional, it cannot be in the range. Also, since we have

$$A\mathbf{a} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 4 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq \mathbf{0},$$

the vector $\mathbf{a}$ is not in the null space of $A$.

---

Next, we consider the vector $\mathbf{b} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$. Again, this is three-dimensional vector, hence $\mathbf{b}$ is not in the range $\mathcal{R}(A)$.
Since we have

$$A\mathbf{b} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0},$$

we see that the vector $\mathbf{b}$ is in the null space $\mathcal{N}(A)$.

---

Finally, we consider the vector $\mathbf{c} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Since it is two-dimensional, the vector $\mathbf{c}$ is not in the null space $\mathcal{N}(A)$.
Let us determine whether $\mathbf{c}$ is in the range $\mathcal{R}(A)$. To do this, we check whether the system $A\mathbf{x} = \mathbf{c}$ is consistent or not. The augmented matrix of the system is

$$[A \mid \mathbf{c}] = \left[\begin{array}{cccc} 1 & 2 & 1 & 1 \\ 3 & 6 & 4 & 1 \end{array}\right].$$

Applying the elementary row operations, we have

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 6 & 4 & 1 \end{bmatrix} \xrightarrow{R_2-3R_1} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix} \xrightarrow{R_1-R_2} \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

and thus it has the general solution

$$\begin{aligned} x_1 &= -2x_2 + 3 \\ x_3 &= -2. \end{aligned}$$

Therefore, the vector **c** is in the range $\mathcal{R}(A)$.

(For example, if $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, then $A\mathbf{x}_0 = \mathbf{c}$.)

In summary, the vector **a** is neither in $\mathcal{N}(A)$ or in $\mathcal{R}(A)$. The vector **b** is in $\mathcal{N}(A)$ but not in $\mathcal{R}(A)$. The vector **c** is not in $\mathcal{N}(A)$ but in $\mathcal{R}(A)$.

## Solution of (b)

We first describe the null space $\mathcal{N}(B)$ of the matrix $B$.

The vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is in the null space $\mathcal{N}(B)$ if and only if $B\mathbf{x} = \mathbf{0}$.

The augmented matrix of $B\mathbf{x} = \mathbf{0}$ is

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ -2 & -2 & -4 & 0 \end{bmatrix}.$$

Applying the elementary row operations, we have

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ -2 & -2 & -4 & 0 \end{bmatrix} \xrightarrow{R_2+2R_1} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and the general solution is

$$x_1 = -x_2 - 2x_3.$$

Therefore, the vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is in the null space $\mathcal{N}(B)$ if and only if $x_1 = -x_2 - 2x_3$.

Thus,

$$\mathcal{N}(B) = \{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathrm{R}^3 \mid x_1 = -x_2 - 2x_3\}$$

$$= \{\mathbf{x} \in \mathrm{R}^3 \mid \mathbf{x} = \begin{bmatrix} -x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix}\}$$

$$= \{\mathbf{x} \in \mathrm{R}^3 \mid \mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\}$$

$$= \mathrm{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Thus, the set

$$S = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a spanning set of the null space $\mathcal{N}(B)$.

Let us verify that $S$ is a linearly independent set. Suppose that we have

$$c_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = \mathbf{0},$$

for some scalars $c_1, c_2$. Then we have

$$\begin{bmatrix} -c_1 - 2c_2 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

and thus $c_1 = c_2 = 0$ and $S$ is linearly independent.

Hence $S$ is a linearly independent spanning set, and thus $S$ is a basis of the null space $\mathcal{N}(B)$.

---

3

**(a)** Let $A = \begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 3 & 1 & 2 \\ 1 & 3 & 1 & 2 \end{bmatrix}$. Find a basis for the range $\mathcal{R}(A)$ of $A$ that consists of columns of $A$.

**(b)** Find the rank and nullity of the matrix $A$ in part (a).

Solution.

### (a) Find a basis for the range $\mathcal{R}(A)$

Note that the range $\mathcal{R}(A)$ is the column space of $A$. Thus,

$$\mathcal{R}(A) = \operatorname{Span}\{A_1, A_2, A_3, A_4\},$$

where $A_i$ is the $i$-th column vector of $A$.

We use the "leading 1" method.

We reduce the matrix $A$ by elementary row operations as follows.

We have

$$A \xrightarrow{R_3 - R_2} \begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The first and the third column contain the leading 1's.

Thus, the set $\{A_1, A_3\}$ is a basis of the range $\mathcal{R}(A)$, which consists of the first and the third column vectors of $A$.

### (b) Find the rank and nullity of the matrix $A$

By part (a), we know that $\{A_1, A_3\}$ is a basis of the range of $A$.

Hence the dimension of the range is 2.

Thus the rank of $A$, which is the dimension of the range $\mathcal{R}(A)$, is 2

Recall the rank-nullity theorem. Since $A$ is a $3 \times 4$ matrix, we have

$$\text{rank of } A + \text{nullity of } A = 4.$$

Since we know that the rank of $A$ is 2, it follows from the rank-nullity theorem that the nullity of $A$ is 2.

---

4 Let

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

be a matrix. Find a basis of the null space of the matrix $A$.

## Solution.

The null space $\mathcal{N}(A)$ of the matrix $A$ is by definition

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathrm{R}^3 \mid A\mathbf{x} = \mathbf{0}\}.$$

In other words, the null space consists of all solutions $\mathbf{x}$ of the matrix equation $A\mathbf{x} = \mathbf{0}$.

So we first determine the solutions of $A\mathbf{x} = \mathbf{0}$ by Gauss-Jordan elimination. The augmented matrix is

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Subtracting $R_1$ from $R_2$, we reduce the augmented matrix to the reduced row echelon form matrix as follows.

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \xrightarrow{R_2-R_1} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus the solution $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ of $A\mathbf{x} = \mathbf{0}$ satisfies $x_1 + x_2 = 0$, or equivalently $x_1 = -x_2$.

Substituting the last equality, we see that solutions are of the form

$$\mathbf{x} = \begin{bmatrix} -x_2 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore the null space is

$$\mathcal{N}(A) = \left\{ \mathbf{x} \in \mathrm{R}^3 \;\middle|\; \mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ for any } x_2, x_3 \in \mathrm{R} \right\}$$

$$= \mathrm{Sp}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Thus, the set $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a spanning set for the null space $\mathcal{N}(A)$.

---

Now, we check that the vectors $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent.

Consider a linear combination

$$a_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{0}.$$

This is equivalent to

$$\begin{bmatrix} -a_1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Hence we must have $a_1 = a_2 = 0$.

Since the linear combination equation has only the zero solution, the vectors $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent.

---

Therefore the set $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a linearly independent spanning set, thus it is a basis for the null space $\mathcal{N}(A)$.

5 Let

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 2 & 3 & 5 \end{bmatrix}.$$

**(a)** Find a matrix $B$ in reduced row echelon form such that $B$ is row equivalent to the matrix $A$.
**(b)** Find a basis for the null space of $A$.
**(c)** Find a basis for the range of $A$ that consists of columns of $A$. For each columns, $A_j$ of $A$ that does not appear in the basis, express $A_j$ as a linear combination of the basis vectors.
**(d)** Exhibit a basis for the row space of $A$.

## Hint.

In part (c), you may use the following theorem.

**Theorem (leading-1 method)**: If $B$ is a matrix in reduced row echelon form that is row equivalent to $A$, then all the column vectors of $A$ whose corresponding columns in $B$ have leading 1's form a basis of the range of $A$.

In part (d), you may use the following theorem.

**Theorem (Row-space method)**: If $B$ is a matrix in (reduced) row echelon form that is row equivalent to $A$, then the nonzero rows of $B$ form a basis for the row space of $A$.

## Solution.

(a) Find a matrix $B$ in reduced roe echelon form such that $B$ is row equivalent to the matrix $A$.

We apply the elementary row operations to the matrix $A$ and obtain

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 2 & 3 & 5 \end{bmatrix} \xrightarrow{\substack{R_2-2R_1 \\ R_3-2R_1}} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1-R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is in reduced row echelon form that is row equivalent to $A$.

Thus we set

$$B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

### (b) Find a basis for the null space of $A$.

The null space $\mathcal{N}(A)$ of the matrix is the set of solutions of the homogeneous system $A\mathbf{x} = \mathbf{0}$. By part (a), the augmented matrix $[A \mid \mathbf{0}]$ is row equivalent to $[B \mid \mathbf{0}]$.

Thus, the solution $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ must satisfy

$$x_1 = -x_3 \text{ and } x_2 = -x_3,$$

where $x_3$ is a free variable.

Thus the solutions are given by

$$\mathbf{x} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

for any number $x_3$.

Hence we have

$$\begin{aligned} \mathcal{N}(A) &= \left\{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \text{ for any } x_3 \in \mathbb{R}\right\} \\ &= \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

From this, we deduce that the set

$$\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

is a basis for the null space of $A$.

### (c) Find a basis for the range of $A$ that consists of columns of $A$. (Longer version)

Let us write

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 2 & 3 & 5 \end{bmatrix} = [A_1, A_2, A_3],$$

where $A_1, A_2, A_3$ are the column vectors of the matrix $A$.
The range $\mathcal{R}(A)$ of $A$ is the same as the column space of $A$.
Thus, it suffices to find the maximal number of linearly independent vectors among the column vectors $A_1, A_2, A_3$.
Consider the linear combination

$$x_1 A_1 + x_2 A_2 + x_3 A_3 = \mathbf{0}. \tag{*}$$

Then this is equivalent to the homogeneous system

$$A\mathbf{x} = \mathbf{0}$$

and we already found the solutions in part (b):

$$x_1 = -x_3 \text{ and } x_2 = -x_3.$$

This tells us the vectors $A_1, A_2, A_3$ are linearly dependent.
For example, $x_1 = -1, x_2 = -1, x_3 = 1$ is a nonzero solution of the system (*).
Thus, we have

$$\mathcal{R}(A) = \operatorname{Span}(A_1, A_2, A_3) = \operatorname{Span}(A_1, A_2).$$

On the other hand, if we consider only $A_1$ and $A_2$, they are linearly independent.
(To see this, you just need to repeat the above argument without $A_3$. This amounts to just ignoring the third columns in the computations.)

Therefore, the set $\{A_1, A_2\}$ is a linearly independent spanning set for the range.
Hence a basis of the range consisting column vectors of $A$ is

$$\{A_1, A_2\}.$$

The only column vector which is not a basis vector is $A_3$.
We already found that $x_1 = -1, x_2 = -1, x_3 = 1$ is a nonzero solution of (*).
Thus we have

$$-A_1 - A_2 + A_3 = \mathbf{0}.$$

Solving this for $A_3$, we obtain the linear combination for $A_3$ of the basis vectors:

$$A_3 = A_1 + A_2.$$

## (c) A shorter solution using the leading 1 method

---

Here is a shorter solution which uses the following theorem.

Theorem (leading-1 method): If $B$ is a matrix in reduced row echelon form that is row equivalent to $A$, then all the column vectors of $A$ whose corresponding columns in $B$ have leading 1's form a basis of the range of $A$.

Looking at the matrix $B$, we see that the first and the second columns of $B$ have leading 1's. Thus the first and the second column vectors of $A$ form a basis for the range of $A$.

(d) Exhibit a basis for the row space of $A$.

We use the following theorem.

Theorem (Row-space method): If $B$ is a matrix in (reduced) row echelon form that is row equivalent to $A$, then the nonzero rows of $B$ form a basis for the row space of $A$.

We have already found such $B$ in part (a), and the first and the second row vectors are nonzero. Thus they form a basis for the row space of $A$.
Hence a basis for the row space of $A$ is

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

### Comment.

The longer solution of part (c) is essentially the proof of Theorem (leading 1 method).
It is good to know where the theorem came from, but when you solve a problem you may forget and just use the theorem like the shorter solution of (c).

---

⑥ Let

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

**(a)** Find a basis for the null space $\mathcal{N}(A)$.
**(b)** Find a basis of the range $\mathcal{R}(A)$.
**(c)** Find a basis of the row space for $A$.
*(The Ohio State University)*

---

⑦ Determine whether each of the following sets is a basis for $\mathbb{R}^3$.

**(a)** $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} \right\}$

**(b)** $S = \left\{ \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \right\}$

**(c)** $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \end{bmatrix} \right\}$

**(d)** $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 6 \end{bmatrix}, \begin{bmatrix} -1 \\ 9 \\ 10 \end{bmatrix} \right\}$

### Definition (A Basis of a Subspace).

A subset $S$ of a vector space $V$ is called a **basis** if

8. $S$ is linearly independent, and
9. $S$ is a spanning set.

### Solution.

Recall that any three linearly independent vectors form a basis of $\mathrm{R}^3$.
(See the post "Three Linearly Independent Vectors in $\mathrm{R}^3$ Form a Basis. Three Vectors Spanning $\mathrm{R}^3$ Form a Basis." for the proof of this fact.)

(a) $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} \right\}$

Let us check that whether $S$ is a linearly independent set.
Consider the linear combination

$$x_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} = \mathbf{0}.$$

This is equivalent to the matrix equation

$$\begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & 1 \\ -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}.$$

To find the solution, consider the augmented matrix.
Applying elementary row operations, we obtain

$$\left[\begin{array}{rrrr} 1 & 2 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & -1 & 4 & 0 \end{array}\right] \xrightarrow{R_3+R_1} \left[\begin{array}{rrrr} 1 & 2 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

$$\xrightarrow[R_3-R_2]{R_1-2R_2} \left[\begin{array}{rrrr} 1 & 0 & -4 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow[R_2-R_3]{R_1+4R_3} \left[\begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

It follows that the solution is $x_1 = x_2 = x_3 = 0$.

Hence $S$ is linearly independent.

As $S$ consists of three linearly independent vectors in $\mathrm{R}^3$, it must be a basis of $\mathrm{R}^3$.

(b) $S = \left\{ \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \right\}$

As in part (a), we determine whether the set $S$ is linearly independent or not by considering the following augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{array}\right] \xrightarrow[R_3 - 7R_1]{R_2 - 4R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & -6 & -12 & 0 \end{array}\right]$$

$$\xrightarrow{-\frac{1}{3}R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -6 & -12 & 0 \end{array}\right] \xrightarrow[R_3 + 6R_2]{R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, the general solution is $x_1 = x_3, x_2 = -2x_3$, where $x_3$ is a free variable.

Hence, in particular, there is a nonzero solution.

So $S$ is linearly dependent, and hence $S$ cannot be a basis for $\mathrm{R}^3$.

(c) $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \end{bmatrix} \right\}$

A quick solution is to note that any basis of $\mathrm{R}^3$ must consist of three vectors. Thus $S$ cannot be a basis as $S$ contains only two vectors.

---

Another solution is to describe the span Span($S$).

Note that a vector $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is in Span($S$) if and only if $\mathbf{v}$ is a linear combination of vectors in $S$.

Equivalently, the vector $\mathbf{v}$ is in Span($S$) if and only if the system

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

is consistent.

Let us consider the augmented matrix and reduce it by elementary row operations.

$$\begin{bmatrix} 1 & 0 & a \\ 1 & 1 & b \\ 2 & 7 & c \end{bmatrix} \xrightarrow[R_3-2R_1]{R_2-R_1} \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b-a \\ 0 & 7 & c-2a \end{bmatrix}$$

$$\xrightarrow{R_3-7R_2} \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b-a \\ 0 & 0 & 5a-7b+c \end{bmatrix}.$$

Note that we obtained the (3, 3)-entry by $c - 2a - 7(b - a) = 5a - 7b + c$.

It follows that the system is consistent if and only if

$$5a - 7b + c = 0.$$

Thus, for example, the vector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is not in Span($S$) as $5 \cdot 1 - 7 \cdot 0 + 0 \neq 0$.

Hence Span($S$) is not $\mathrm{R}^3$, and we conclude that $S$ is not a basis.

(d) $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 6 \end{bmatrix}, \begin{bmatrix} -1 \\ 9 \\ 10 \end{bmatrix} \right\}$

The set $S$ contains four 3-dimensional vectors. Hence $S$ is linearly dependent, and thus $S$ is not a basis.

---

**10** Suppose that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -4 \\ 0 \\ -3 \\ -2 \\ 1 \end{bmatrix}$$

are a basis vectors for the null space of a $4 \times 5$ matrix $A$. Find a vector $\mathbf{x}$ such that $\mathbf{x} \neq 0$, $\mathbf{x} \neq \mathbf{v}_1$, $\mathbf{x} \neq \mathbf{v}_2$ and $A\mathbf{x} = \mathbf{0}$.

(*Stanford University*)

## Solution.

We are asked to find a vector $\mathbf{x}$ in the null space of $A$, which are not $\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2$.

Recall that the null space is a vector space. Thus, any linear combination of vectors in the null space is still in the null space.

Since $\mathbf{v}_1, \mathbf{v}_2$ are basis of the null space of $A$, they are in particular vectors in the null space of $A$. Thus, for example,

$$\mathbf{x} = 2\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

is an element of the null space and it is not equal to $\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2$.

Another example is

$$\mathbf{x} = \mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} -6 \\ 1 \\ -3 \\ -2 \\ 1 \end{bmatrix}.$$

In general, you can prove that any vector of the form

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2,$$

where $c_1, c_2$ are scalars such that $(c_1, c_2) \neq (0,0), (1,0), (0,1)$, satisfied the required conditions.

---

⑪ Let $V$ be the following subspace of the $4$-dimensional vector space $\mathbb{R}^4$.

$$V := \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; x_1 - x_2 + x_3 - x_4 = 0 \right\}.$$

Find a basis of the subspace $V$ and its dimension.

**Solution.**

Any vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ in $V$ satisfies $x_1 - x_2 + x_3 - x_4 = 0$, or equivalently $x_1 = x_2 - x_3 + x_4$.

Thus we have

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_2 - x_3 + x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$= x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The above computation shows that any vector $\mathbf{x}$ in $V$ can be written as a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$.
Hence the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a spanning set for the subspace $V$.

---

We claim that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a linearly independent set.
Consider

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + a_3\mathbf{u}_3 = \mathbf{0}.$$

We show that this equation has only the zero solution $a_1 = a_2 = a_3 = 0$.
The equation (*) can be written as

$$\begin{bmatrix} a_1 - a_2 + a_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \mathbf{0}.$$

Comparing entries, we obtain $a_1 = a_2 = a_3 = 0$.
Thus the equation (*) has only the zero solution and hence the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent.

---

Therefore, the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is linearly independent spanning set for $V$, thus it is a basis for the subspace $V$.
Since the basis consists of 3 vectors, the dimension of the subspace $V$ is 3.

---

In summary, we found a basis

$$\left\{ \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

for the subspace $V$ and the dimension of $V$ is 3.

---

⑫

A **hyperplane** in $n$-dimensional vector space $\mathbb{R}^n$ is defined to be the set of vectors $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ satisfying the linear equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \dots, a_n$ (at least one of $a_1, a_2, \dots, a_n$ is nonzero) and $b$ are real numbers. Here at least one of $a_1, a_2, \dots, a_n$ is nonzero. Consider the hyperplane $P$ in $\mathbb{R}^n$ described by the linear equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0,$$

where $a_1, a_2, \dots, a_n$ are some fixed real numbers and not all of these are zero.
(The constant term $b$ is zero.) Then prove that the hyperplane $P$ is a subspace of $R^n$ of dimension $n - 1$.

## Proof.

### The hyperplane is a subspace.

In a set theoretical notation, the hyperplane is give by

$$P = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathrm{R}^n \;\middle|\; a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \right\}$$

The defining relation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

can be written as

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

Let $A$ be the $1 \times n$ matrix $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$.

Then the above equation is simply $A\mathbf{x} = 0$, and the hyperplane $P$ is described by

$$P = \{\mathbf{x} \in \mathrm{R}^n \mid A\mathbf{x} = 0\},$$

and this is exactly the definition of the null space of $A$. Namely, we have

$$P = \mathcal{N}(A).$$

Since every null space of a matrix is a subspace, it follows that the hyperplane $P$ is a subspace of $\mathrm{R}^n$.

The dimension of the hyperplane is $n - 1$.

Because $P = \mathcal{N}(A)$, the dimension of $P$ is the nullity of the matrix $A$.

Since not all of $a_i$'s are zero, the rank of the matrix $A$ is 1. Then by the rank-nullity theorem, we have

$$\text{rank of } A + \text{nullity of } A = n.$$

Hence the nullity of $A$ is $n - 1$.

---

⑬ Let $A$ be a real $7 \times 3$ matrix such that its null space is spanned by the vectors

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Then find the rank of the matrix $A$.
(*Purdue University*)

## Solution.

We first determine the nullity of $A$ and deduce the rank of $A$ by the rank-nullity theorem.

The null space $\mathcal{N}(A)$ of the matrix $A$ is spanned by

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Let us find a basis of the null space $\mathcal{N}(A)$ among these vectors.

We use the “leading 1 method”.

Form the matrix whose column vectors are these three vectors and we apply elementary row operations as follows.

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2}$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is in reduced row echelon form and the first and second column contain the leading 1’s.

Therefore, the first two vectors

$$\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}$$

form a basis of the null space $\mathcal{N}(A)$.

Hence the nullity, which is the dimension of $\mathcal{N}(A)$, is 2.

Since the size of the matrix $A$ is $7 \times 3$, the rank-nullity theorem gives

$$3 = \text{nullity of } A + \text{rank of } A.$$

Thus, the rank of $A$ is 1.

---

14 Let $A = \begin{bmatrix}4 & 1\\3 & 2\end{bmatrix}$ and consider the following subset $V$ of the 2-dimensional vector space $\mathbb{R}^2$.

$$V = \{\mathbf{x} \in \mathbb{R}^2 \mid A\mathbf{x} = 5\mathbf{x}\}.$$

**(a)** Prove that the subset $V$ is a subspace of $\mathbb{R}^2$.

**(b)** Find a basis for $V$ and determine the dimension of $V$.

## Proof.

### (a) $V$ is a subspace of $\mathrm{R}^2$

Note that the equality $A\mathbf{x} = 5\mathbf{x}$ can be written as

$$(A - 5I)\mathbf{x} = \mathbf{0},$$

where $I$ is the $2 \times 2$ matrix.

If we define $B = A - 5I$, then the subset $V$ becomes

$$\begin{aligned} V &= \{\mathbf{x} \in \mathrm{R}^2 \mid B\mathbf{x} = \mathbf{0}\} \\ &= \mathcal{N}(B), \end{aligned}$$

the null space of the matrix $B$.

Since any null space is a vector space, this shows that $V$ is a subspace of $\mathrm{R}^2$.

### (b) A basis for $V$ and the dimension of $V$

From part (a), we have obtained that

$$V = \mathcal{N}(B),$$

where

$$B = A - 5I = \begin{bmatrix}-1 & 1\\3 & -3\end{bmatrix}.$$

Thus, $V$ is a set of solutions of the system $B\mathbf{x} = \mathbf{0}$.

The augmented matrix for this system is reduced as follows.

$$\begin{bmatrix} -1 & 1 & 0 \\ 3 & -3 & 0 \end{bmatrix} \xrightarrow{R_2+3R_1} \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-R_1} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

From this we see that the general solution $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ satisfies

$$x_1 = x_2,$$

and thus the general solution is

$$\mathbf{x} = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Therefore we have

$$\begin{aligned} V &= \left\{ \mathbf{x} \in \mathrm{R}^2 \;\middle|\; \mathbf{x} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for any } x_2 \in \mathrm{R} \right\} \\ &= \operatorname{Span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Thus the set

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

is a spanning set of $V$ and it is linearly independent as it consists of only a single vector.

Therefore

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for the subspace $V$, and thus the dimension of $V$ is 1.

(Recall that the dimension of a vector space is the number of vectors in a basis for the vector space.)

---

**15** Let $V$ be a subset of $\mathbb{R}^4$ consisting of vectors that are perpendicular to vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, where

$$\mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}.$$

Namely,

$$V = \{\mathbf{x} \in \mathbb{R}^4 \mid \mathbf{a}^\mathrm{T}\mathbf{x} = 0, \mathbf{b}^\mathrm{T}\mathbf{x} = 0, \text{ and } \mathbf{c}^\mathrm{T}\mathbf{x} = 0\}.$$

**(a)** Prove that $V$ is a subspace of $\mathbb{R}^4$.

**(b)** Find a basis of $V$.

**(c)** Determine the dimension of $V$.

Proof.

### (a) Prove that $V$ is a subspace of $\mathbb{R}^4$.

Observe that the conditions

$$\mathbf{a}^{\mathrm{T}}\mathbf{x} = 0, \mathbf{b}^{\mathrm{T}}\mathbf{x} = 0, \text{ and } \mathbf{c}^{\mathrm{T}}\mathbf{x} = 0$$

can be combined into the following matrix equation

$$A\mathbf{x} = \mathbf{0},$$

where

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

and $\mathbf{0}$ is the three dimensional zero vector.

Note that the rows of the matrix $A$ are $\mathbf{a}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}}$, and $\mathbf{c}^{\mathrm{T}}$.

It follows that the subset $V$ is the null space $\mathcal{N}(A)$ of the matrix $A$.

Being the null space, $V = \mathcal{N}(A)$ is a subspace of $\mathbb{R}^4$.

(See the post "The Null Space (the Kernel) of a Matrix is a Subspace of $\mathbb{R}^n$".)

### (b) Find a basis of $V$.

In the proof of Part (a), we saw that $V = \mathcal{N}(A)$.

To find a basis, we determine the solutions of $A\mathbf{x} = \mathbf{0}$.

Applying elementary row operations to the augmented matrix, we see that

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 - R_2} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

It follows that the general solution is given by

$$x_1 = -x_3, x_2 = x_3.$$

The vector form solution is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_3 \\ x_3 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Hence we have

$$\begin{aligned} V &= \mathcal{N}(A) \\ &= \left\{ \mathbf{x} \in \mathbb{R}^4 \;\middle|\; \mathbf{x} = x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \text{ where } x_3, x_4 \in \mathbb{R} \right\} \\ &= \operatorname{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Let $B := \left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Then we just showed that $B$ is a spanning set for $V$.

It is straightforward to see that $B$ is linearly independent.

Hence $B$ is a basis for $V$.

(c) Determine the dimension of $V$.

As the basis $B$ for $V$ that we obtained in Part (b) consists of two vectors, the dimension of the subspace $V$ is $\dim(V) = 2$.

---

(16) Find a basis for the subspace $W$ of all vectors in $\mathbb{R}^4$ which are perpendicular to the columns of the matrix

$$A = \begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix}.$$

(*Harvard University Exam*)

Hint.

(17) Show that $W = \mathcal{N}(A^{\mathrm{T}})$.

(18) Find a basis of $\mathcal{N}(A^{\mathrm{T}})$ by reducing the matrix $A^{\mathrm{T}}$.

Solution.

Let us write $A = [A_1\ A_2\ A_3\ A_4]$, where $A_i$ is the $i$-th column vector of $A$ for $i = 1, 2, 3, 4$.

First we claim that a vector $\mathbf{x} \in \mathbb{R}^4$ is perpendicular to all column vectors $A_i$ if and only if $\mathbf{x} \in \mathcal{N}(A^{\mathrm{T}})$.

To see this, we compute

$$A^{\mathrm{T}}\mathbf{x} = \begin{bmatrix} A_1^{\mathrm{T}} \\ A_2^{\mathrm{T}} \\ A_3^{\mathrm{T}} \\ A_4^{\mathrm{T}} \end{bmatrix} \mathbf{x} = \begin{bmatrix} A_1^{\mathrm{T}}\mathbf{x} \\ A_2^{\mathrm{T}}\mathbf{x} \\ A_3^{\mathrm{T}}\mathbf{x} \\ A_4^{\mathrm{T}}\mathbf{x} \end{bmatrix}.$$

From this equality the claim follows immediately.

So we proved that $\mathcal{N}(A^{\mathrm{T}}) = W$. From this, we see that $W$ is actually a subspace of $\mathbb{R}^4$.

---

Thus, we need to find a basis for the null space of the transpose $A^{\mathrm{T}}$.

We apply elementary row operations to $A^{\mathrm{T}}$ and obtain a reduced row echelon form

$$A^{\mathrm{T}} \to \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The last two columns correspond to two free variables. Let $s$ and $t$ be free variable.

Then $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathcal{N}(A^{\mathrm{T}})$ if and only if $\mathbf{x}$ satisfies

$$\begin{aligned} x_1 &= s + 2t \\ x_2 &= -2s - 3t \\ x_3 &= s \\ x_4 &= t, \end{aligned}$$

equivalently

$$\mathbf{x} = s \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore a basis of $W = \mathcal{N}(A^{\mathrm{T}})$ is

$$\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}.$$

---

⑲ Let $V$ be the vector space over $\mathbb{R}$ of all real $2 \times 2$ matrices. Let $W$ be the subset of $V$ consisting of all symmetric matrices.
**(a)** Prove that $W$ is a subspace of $V$.
**(b)** Find a basis of $W$.
**(c)** Determine the dimension of $W$.

Proof.

Recall that $A$ is **symmetric** if $A^{\mathrm{T}}=A$.

(a) Prove that $W$ is a subspace of $V$.

We verify the following subspace criteria:

**Subspace Criteria**.

⑳ The zero vector of $V$ is in $W$.

㉑ For any $A, B \in W$, the sum $A+B \in W$.

㉒ For any $A \in W$ and $r \in \mathbb{R}$, the scalar product $rA \in W$.

The zero vector in $V$ is the $2 \times 2$ zero matrix $O$.
It is clear that $O^{\mathrm{T}}=O$, and hence $O$ is symmetric.
Thus $O \in W$ and condition 1 is met.

Let $A, B$ be arbitrary elements in $W$.
That is, $A$ and $B$ are symmetric matrices.
We show that the sum $A+B$ is also symmetric.
We have

$$(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}=A+B.$$

The second equality follows as $A, B$ are symmetric.
Hence $A+B$ is symmetric and $A+B \in W$.
Condition 2 is met.

To check condition 3, let $A \in W$ and $r \in \mathbb{R}$.
We have

$$(rA)^{\mathrm{T}}=rA^{\mathrm{T}}=rA,$$

where the second equality follows since $A$ is symmetric.
This implies that $rA$ is symmetric, and hence $rA \in W$.
So condition 3 is met, and we conclude that $W$ is a subspace of $V$ by subspace criteria.

### (b) Find a basis of $W$.

Let

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

be an arbitrary element in the subspace $W$.

Then since $A^{\mathrm{T}}=A$, we have

$$\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

This implies that $a_{12}=a_{21}$, and hence

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \\ &= a_{11}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}+a_{12}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}+a_{22}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Let $B=\{v_1, v_2, v_3\}$, where $v_1, v_2, v_3$ are $2\times 2$ matrices appearing in the above linear combination of $A$.

Note that these matrices are symmetric.

Hence we showed that any element in $W$ is a linear combination of matrices in $B$.

Thus $B$ is a spanning set for the subspace $W$.

---

We show that $B$ is linearly independent.

Suppose that we have

$$c_1v_1+c_2v_2+c_3v_3=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then it follows that

$$\begin{bmatrix} c_1 & c_2 \\ c_2 & c_3 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus $c_1=c_2=c_3=$ and the set $B$ is linearly independent.

As $B$ is a linearly independent spanning set, we conclude that $B$ is a basis for the subspace $W$.

### (c) Determine the dimension of $W$.

Recall that the dimension of a subspace is the number of vectors in a basis of the subspace.

In part (b), we found that $B=\{v_1, v_2, v_3\}$ is a basis for the subspace $W$.

As $B$ consists of three vectors, the dimension of $W$ is $3$.

## Related Question (Skew-Symmetric Matrices)

A matrix $A$ is called **skew-symmetric** if $A^{\mathrm{T}}=-A$.

**Problem.**

Let $V$ be the vector space of all $2 \times 2$ matrices.

Let $W$ be a subset of $V$ consisting of all $2 \times 2$ skew-symmetric matrices.

Prove that $W$ is a subspace of $V$ and also find a basis and dimension of $W$.

---

23 Suppose that $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ are linearly dependent $n$-dimensional real vectors. For any vector $\mathbf{v}_{r+1} \in \mathbb{R}^n$, determine whether the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}$ are linearly independent or linearly dependent.

## Solution.

We claim that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}$ are linearly dependent.

Since the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ are linearly dependent, there exist scalars (real numbers) $a_1, a_2, \dots, a_r$ such that

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = \mathbf{0}$$

and not all of $a_1, \dots, a_r$ are zero, that is, $(a_1, \dots, a_r) \neq (0, \dots, 0)$.

Consider the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_r\mathbf{v}_r + x_{r+1}\mathbf{v}_{r+1} = \mathbf{0}.$$

If this equation has a nonzero solution $(x_1, \dots, x_r, x_{r+1})$, then the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{r+1}$ are linearly dependent.

---

In fact,

$$(x_1, x_2, \dots, x_r, x_{r+1}) = (a_1, a_2, \dots, a_r, 0)$$

is a nonzero solution of the above equation.

To see this, first note that since not all of $a_1, a_2, \dots, a_r$ are zero, we have

$$(a_1, a_2, \dots, a_r, 0) \neq (0, 0, \dots, 0, 0).$$

Plug these values in the equation, we have

$$\begin{aligned} &a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r + 0\mathbf{v}_{r+1} \\ &= a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = \mathbf{0} \text{ by } (*). \end{aligned}$$

Therefore, we conclude that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}$ are linearly dependent.

---

24 Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a set of three-dimensional vectors in $\mathbb{R}^3$.

**(a)** Prove that if the set $B$ is linearly independent, then $B$ is a basis of the vector space $\mathbb{R}^3$.

**(b)** Prove that if the set $B$ spans $\mathbb{R}^3$, then $B$ is a basis of $\mathbb{R}^3$.

## Definition (A Basis of a Subspace).

A subset $B$ of a vector space $V$ is called a **basis** if $B$ is linearly independent spanning set.

## Proof.

### (a) Prove that if the set $B$ is linearly independent, then $B$ is a basis of the vector space $\mathrm{R}^3$.

To show that $B$ is a basis, we need only prove that $B$ is a spanning set of $\mathrm{R}^3$ as we know that $B$ is linearly independent.
Let $\mathbf{b} \in \mathrm{R}^3$ be an arbitrary vector.
We prove that there exist $x_1, x_2, x_3$ such that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}.$$

This is equivalent to having a solution $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ to the matrix equation

$$A\mathbf{x} = \mathbf{b},$$

where

$$A = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$$

is the $3 \times 3$ matrix whose column vectors are $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Since the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent, the matrix $A$ is nonsingular.
It follows that the equation (*) has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.
Hence $\mathbf{b}$ is a linear combination of the vectors in $B$.
This means that $B$ is a spanning set of $\mathrm{R}^3$, hence $B$ is a basis.

### (b) Prove that if the set $B$ spans $\mathrm{R}^3$, then $B$ is a basis of $\mathrm{R}^3$.

As we know that $B$ spans $\mathrm{R}^3$, it suffices to show that $B$ is linearly independent.
Note that the assumption that $B$ is a spanning set of $\mathrm{R}^3$ means that any vector $\mathbf{b}$ in $\mathrm{R}^3$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$: there exist $c_1, c_2, c_3$ such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{b}.$$

Equivalently, for any $\mathbf{b} \in \mathrm{R}^3$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

Let $A = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ as before.
It follows that the equation

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

has a solution $\mathbf{x} = \mathbf{u}_1$.
Similarly the equations

$$A\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad A\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

have solutions $\mathbf{u}_2, \mathbf{u}_3$, respectively.

---

Define the 3 × 3 matrix $A'$ by

$$A' = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3].$$

Then it follows that

$$\begin{aligned} AA' &= A[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] \\ &= [A\mathbf{u}_1, A\mathbf{u}_2, A\mathbf{u}_3] \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

---

As the identity matrix is nonsingular, the product $AA'$ is nonsingular.

Thus, the matrix $A$ is nonsingular as well.

This implies that the column vectors of $A$ are linearly independent.

Hence the set $B$ is linearly independent and we conclude that $B$ is a basis of $\mathrm{R}^3$.

25 Let $V$ be a subspace of $\mathbb{R}^n$. Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is a spanning set for $V$. Prove that any set of $m+1$ or more vectors in $V$ is linearly dependent.

We give two proofs. The essential ideas behind those two proofs are the same.

The second proofs uses matrix equations to organize clutter of a system of linear equations.

## Proof 1.

Consider the set $T = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ of $k$ vectors in the subspace $V$, where $k \geq m + 1$.

We prove that the set $T$ is linearly dependent.

Consider the linear combination

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_k\mathbf{u}_k = \mathbf{0},$$

where $\mathbf{0}$ is the zero vector of $\mathrm{R}^n$.

Our goal is to find a nonzero solution $x_1, x_2, \dots, x_n$ of this equation.

---

As $S$ is a spanning set for $V$, each vector $\mathbf{v}_i$ in $T$ can be written as a linear combination of vectors in $S$.

So we have

$$\begin{aligned}\mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{m1}\mathbf{v}_m\\ \mathbf{u}_2 &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{m2}\mathbf{v}_m\\ \vdots \quad &\quad \vdots\\ \mathbf{u}_k &= a_{1k}\mathbf{v}_1 + a_{2k}\mathbf{v}_2 + \cdots + a_{mk}\mathbf{v}_m.\end{aligned}$$

Here $a_{ij} \in \mathrm{R}$ for $1 \leq i \leq m$, $1 \leq j \leq k$.

Using these linear combination, the equation (*) can be written as

$$\begin{aligned}x_1(a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{m1}\mathbf{v}_m) + x_2(a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{m2}\mathbf{v}_m)\\ + \cdots + x_k(a_{1k}\mathbf{v}_1 + a_{2k}\mathbf{v}_2 + \cdots + a_{mk}\mathbf{v}_m) = \mathbf{0}.\end{aligned}$$

Rearranging terms, we obtain

$$\begin{aligned}(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k)\mathbf{v}_1 + (a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k)\mathbf{v}_2\\ + \cdots + (a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mk}x_k)\mathbf{v}_m = \mathbf{0}.\end{aligned}$$

---

Now it suffices to find a nonzero $x_1, \ldots x_k$ such that the coefficients of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ in (**) are all zero. Thus, we need to find a nonzero solution of the following system.

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k &= 0\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k &= 0\\ &\vdots\\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mk}x_k &= 0.\end{aligned}$$

Observe that this is a $m \times k$ homogeneous system and as $m + 1 \leq k$, there are more variables than equations.
Hence this homogeneous system has infinitely many solutions, in particular, it has a nonzero solution $x_1, \ldots, x_k$.
Hence the equation (**) and, equivalently, the equation (*) has a nonzero solution $x_1, \ldots, x_k$.
Therefore the set $T$ is linearly dependent. This completes the proof.

## Proof 2.

Let $T = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ be the set of $k$ vectors in the subspace $V$, where $k \geq m + 1$.
We want to show that $T$ is linearly dependent.
This is equivalent to finding a nonzero solution $x_1, \ldots, x_k$ of the equation

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_k\mathbf{u}_k = \mathbf{0}.$$

The equation can be written as

$$U\mathbf{x} = \mathbf{0},$$

where

$$U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k] \text{ and } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}.$$

Note that $U$ is an $n \times k$ matrix.

---

Since the subspace $V$ is spanned by the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$, we have

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{m1}\mathbf{v}_m \\ \mathbf{u}_2 &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{m2}\mathbf{v}_m \\ \vdots \quad & \quad \vdots \\ \mathbf{u}_k &= a_{1k}\mathbf{v}_1 + a_{2k}\mathbf{v}_2 + \cdots + a_{mk}\mathbf{v}_m. \end{aligned}$$

Let $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m]$ be the $n \times m$ matrix whose column vectors are $\mathbf{v}_1, \dots, \mathbf{v}_m$.
Then the above linear combinations can be expressed as

$$V \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} = \mathbf{u}_1, V \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} = \mathbf{u}_2, \dots, V \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} = \mathbf{u}_k.$$

Combining these into one matrix equation, we obtain

$$VA = U,$$

where

Every Basis of a Subspace Has the Same Number of Vectors

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}.$$

---

Thus, the equation (***) becomes

$$VA\mathbf{x} = \mathbf{0}.$$

Note that $A$ is an $m \times k$ matrix with $k > m$.
So the homogeneous system $A\mathbf{x} = \mathbf{0}$ has a nonzero solution $\mathbf{x} = \mathbf{x}_0$ since there are more variables than equations.

Hence we have $VA\mathbf{x}_0 = V\mathbf{0} = \mathbf{0}$, and the equation (***) has a nonzero solution $\mathbf{x}_0$.
Thus we conclude that the set $T$ is linearly dependent, which completes the proof.

## Related Question.

As an application of this problem, you may want to prove the following problem.

**Problem.**

Let $V$ be a subspace of $\mathrm{R}^n$.

Suppose that $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a basis of the subspace $V$.

Prove that every basis of $V$ consists of $k$ vectors in $V$.

---

26 Let $V$ be a subspace of $\mathbb{R}^n$. Suppose that $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a basis of the subspace $V$. Prove that every basis of $V$ consists of $k$ vectors in $V$.

## Hint.

You may use the following fact:

**Fact.**

If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is a spanning set of a subspace $V$ of $\mathrm{R}^n$, then any set of $m + 1$ or more vectors of $V$ is linearly dependent.

For a proof of this fact, see the post ↴

If there are More Vectors Than a Spanning Set, then Vectors are Linearly Dependent

## Proof.

Let $B' = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_l\}$ be an arbitrary basis of the subspace $V$.

Our goal is to show that $l = k$.

---

As $B$ is a basis, it is a spanning set for $V$ consisting of $k$ vectors.

By the fact stated above, a set of $k + 1$ or more vectors of $V$ must be linearly dependent.

Since $B'$ is a basis, it is linearly independent.

It follows that $l \leq k$.

---

We now change the roles of $B$ and $B'$.

As $B'$ is a basis, it is a spanning set for $V$ consisting of $l$ vectors.

So it follows from Fact that a set of $l + 1$ or more vectors must be linearly dependent.

Since $B$ is a basis, it is linearly independent.

Hence $k \leq l$.

---

Therefore we have $l \leq k$ and $k \leq l$, and it yields that $l = k$, as required.

**27** Prove the followings.
**(a)** $\mathcal{N}(A) = \mathcal{N}(A^{\mathrm{T}}A)$.
**(b)** $\operatorname{rank}(A) = \operatorname{rank}(A^{\mathrm{T}}A)$.

## Hint.

For part (b), use the rank-nullity theorem and the result from (a).

The rank-nullity theorem says that for a $m \times n$ matrix,

$$\text{rank of } A + \text{ nullity of } A = n.$$

## Proof.

### (a)$\mathcal{N}(A) = \mathcal{N}(A^{\mathrm{T}}A)$.

Show $\mathcal{N}(A) \subset \mathcal{N}(A^{\mathrm{T}}A)$

Consider any $\mathbf{x} \in \mathcal{N}(A)$. Then we have $A\mathbf{x} = \mathbf{0}$. Multiplying it by $A^{\mathrm{T}}$ from the left, we obtain

$$A^{\mathrm{T}}A\mathbf{x} = A^{\mathrm{T}}\mathbf{0} = \mathbf{0}.$$

Thus $\mathbf{x} \in \mathcal{N}(A^{\mathrm{T}}A)$, and hence $\mathcal{N}(A) \subset \mathcal{N}(A^{\mathrm{T}}A)$.

Show $\mathcal{N}(A) \supset \mathcal{N}(A^{\mathrm{T}}A)$

On the other hand, let $\mathbf{x} \in \mathcal{N}(A^{\mathrm{T}}A)$. Thus we have

$$A^{\mathrm{T}}A\mathbf{x} = \mathbf{0}.$$

Multiplying it by $\mathbf{x}^{\mathrm{T}}$ from the left, we obtain

$$\mathbf{x}^{\mathrm{T}}A^{\mathrm{T}}A\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{0} = \mathbf{0}.$$

This implies that we have

$$\mathbf{0} = (A\mathbf{x})^{\mathrm{T}}(A\mathbf{x}) = ||A\mathbf{x}||^2$$

and the length of the vector $A\mathbf{x}$ is zero, thus the vector $A\mathbf{x} = \mathbf{0}$. Hence $\mathbf{x} \in \mathcal{N}(A)$, and we obtain $\mathcal{N}(A) \supset \mathcal{N}(A^{\mathrm{T}}A)$.

### (b) $\operatorname{rank}(A) = \operatorname{rank}(A^{\mathrm{T}}A)$

We use the rank-nullity theorem and obtain

$$\operatorname{rank}(A) = n - \dim(\mathcal{N}(A)) = n - \dim(\mathcal{N}(A^{\mathrm{T}}A)) = \operatorname{rank}(A^{\mathrm{T}}A).$$

(Note that the size of the matrix $A^{\mathrm{T}}A$ is $n \times n$.)

---

**28** Let $A$ be an $m \times n$ matrix and $B$ be an $n \times l$ matrix. Then prove the followings.
**(a)** $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$.

**(b)** If the matrix $B$ is nonsingular, then $\mathrm{rank}(AB) = \mathrm{rank}(A)$.

## Hint.

The rank of an $m \times n$ matrix $M$ is the dimension of the range $\mathcal{R}(M)$ of the matrix $M$.
The range of the matrix $M$ is

$$\mathcal{R}(M) = \{\mathbf{y} \in \mathrm{R}^m \mid \mathbf{y} = M\mathbf{x} \text{ for some } \mathbf{x} \in \mathrm{R}^n\}.$$

## Proof.

### (a) rank($AB$) ≤ rank($A$).

Recall that the rank of a matrix $M$ is the dimension of the range $\mathcal{R}(M)$ of the matrix $M$.
So we have

$$\mathrm{rank}(AB) = \dim(\mathcal{R}(AB)), \quad \mathrm{rank}(A) = \dim(\mathcal{R}(A)).$$

In general, if a vector space $V$ is a subset of a vector space $W$, then we have

$$\dim(V) \leq \dim(W).$$

Thus, it suffices to show that the vector space $\mathcal{R}(AB)$ is a subset of the vector space $\mathcal{R}(A)$.

Consider any vector $\mathbf{y} \in \mathcal{R}(AB)$. Then there exists a vector $\mathbf{x} \in R^l$ such that $\mathbf{y} = (AB)\mathbf{x}$ by the definition of the range.
Let $\mathbf{z} = B\mathbf{x} \in \mathrm{R}^n$.

Then we have

$$\mathbf{y} = A(B\mathbf{x}) = A\mathbf{z}$$

and thus the vector $\mathbf{y}$ is in $\mathcal{R}(A)$. Thus $\mathcal{R}(AB)$ is a subset of $\mathcal{R}(A)$ and we have

$$\mathrm{rank}(AB) = \dim(\mathcal{R}(AB)) \leq \dim(\mathcal{R}(A)) = \mathrm{rank}(A)$$

as required.

### (b) If the matrix $B$ is nonsingular, then rank($AB$) = rank($A$).

Since the matrix $B$ is nonsingular, it is invertible. Thus the inverse matrix $B^{-1}$ exists. We apply part (a) with the matrices $AB$ and $B^{-1}$, instead of $A$ and $B$. Then we have

$$\mathrm{rank}((AB)B^{-1}) \leq \mathrm{rank}(AB)$$

from (a).

Combining this with the result of (a), we have

$$\mathrm{rank}(A) = \mathrm{rank}((AB)B^{-1}) \leq \mathrm{rank}(AB) \leq \mathrm{rank}(A).$$

Therefore all the inequalities are in fact equalities, and hence we have

$$\mathrm{rank}(AB) = \mathrm{rank}(A).$$

㉙ Let $V$ and $W$ be subspaces of $\mathbb{R}^n$ such that $V \cap W = \{\mathbf{0}\}$ and $\dim(V) + \dim(W) = n$.
**(a)** If $\mathbf{v} + \mathbf{w} = \mathbf{0}$, where $\mathbf{v} \in V$ and $\mathbf{w} \in W$, then show that $\mathbf{v} = \mathbf{0}$ and $\mathbf{w} = \mathbf{0}$.
**(b)** If $B_1$ is a basis for the subspace $V$ and $B_2$ is a basis for the subspace $W$, then show that the union $B_1 \cup B_2$ is a basis for $R^n$.
**(c)** If $\mathbf{x}$ is in $\mathbb{R}^n$, then show that $\mathbf{x}$ can be written in the form $\mathbf{x} = \mathbf{v} + \mathbf{w}$, where $\mathbf{v} \in V$ and $\mathbf{w} \in W$.
**(d)** Show that the representation obtained in part (c) is unique.

### Hint.

Try to show that the vector $\mathbf{v}$ is in both $V$ and $W$, hence in $V \cap W$.

### Proof.

#### (a) Show that $\mathbf{v} = \mathbf{0}$ and $\mathbf{w}=\mathbf{0}

Suppose that we have $\mathbf{v} + \mathbf{w} = \mathbf{0}$, where $\mathbf{v} \in V$ and $\mathbf{w} \in W$. Then we have $\mathbf{v} = -\mathbf{w}$. Note that the left hand side is in $V$ and the right hand side is in $W$. Thus the vector $\mathbf{v}$ is in both $V$ and $W$, that is $\mathbf{v} \in V \cap W = \{\mathbf{0}\}$. Hence $\mathbf{v} = \mathbf{0}$, and it follows that $\mathbf{w}$ is also the zero vector.

#### (b) Show that the union $B_1 \cup B_2$ is a basis for $R^n$

Let $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ and $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_l\}$ be bases for $V$ and $W$, respectively.
Note that $\dim(V) = k$, $\dim(W) = l$, and $k + l = n$.

We claim that the set $B_1 \cup B_2$ is linearly independent.
Suppose that we have a linear combination of vectors in $B_1 \cup B_2$

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_l\mathbf{w}_l = \mathbf{0}.$$

Then observe that the vector $\mathbf{v} := c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ is in $V$ and the vector $\mathbf{w} := d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_l\mathbf{w}_l$ is in $W$. Since their sum $\mathbf{v} + \mathbf{w} = \mathbf{0}$, by part (a), we must have $\mathbf{v} = \mathbf{0}$ and $\mathbf{w} = \mathbf{0}$.

Since $B_1$ is a basis for $V$, the vectors $\mathbf{v}_i$ are linearly independent. Thus

$$\mathbf{0} = \mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

implies that the coefficients $c_i$ are all zero.

Similarly, since $B_2$ is a basis for $W$,

$$\mathbf{0} = \mathbf{w} = d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_l\mathbf{w}_l$$

implies that we have $d_i = 0$.

Therefore the coefficient of the linear combination (*) is all zero. Thus the vectors $\mathbf{v_1}, \dots, \mathbf{v_k}, \mathbf{w}_1, \dots, \mathbf{w}_l$ are linearly independent.
Since $B_1 \cup B_2$ consists of $n$ linearly independent vectors in the $n$ dimensional vector space $\mathrm{R}^n$, they must be a basis.

#### (c) Show that $\mathbf{x}$ can be written in the form $\mathbf{x} = \mathbf{v} + \mathbf{w}$

For any vector $\mathbf{x} \in \mathrm{R}^n$, since $B_1 \cup B_2$ is a basis we have

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_l\mathbf{w}_l$$

for some $c_1, \ldots, c_k, d_1, \ldots, d_l \in \mathrm{R}$.

Let

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

and

$$\mathbf{w} = d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_l\mathbf{w}_l.$$

Then we have $\mathbf{v} \in V$, $\mathbf{w} \in W$, and $\mathbf{x} = \mathbf{v} + \mathbf{w}$ as required.

> (d) Show that the representation obtained in part (c) is unique.

Suppose that we have two representations for $\mathbf{x}$, that is, we have $\mathbf{x} = \mathbf{v}' + \mathbf{w}'$ in addition to the $\mathbf{x} = \mathbf{v} + \mathbf{w}$, where $\mathbf{v}, \mathbf{v}' \in V$ and $\mathbf{w}, \mathbf{w}' \in W$.

Then we have $\mathbf{v} + \mathbf{w} = \mathbf{x} = \mathbf{v}' + \mathbf{w}'$, or equivalently,

$$\mathbf{v} - \mathbf{v}' = \mathbf{w}' - \mathbf{w}.$$

Observe that the left hand side is in $V$ and the right hand side is in $W$ thus we have

$$\mathbf{v} - \mathbf{v}', \mathbf{w}' - \mathbf{w} \in V \cap W = \{\mathbf{0}\}.$$

Thus we have $\mathbf{v} = \mathbf{v}'$ and $\mathbf{w} = \mathbf{w}'$ and we conclude that the representations are actually identical, hence we proved the uniqueness.

---

30 Let $A$ be an $m \times n$ matrix. Prove that the rank of $A$ is the same as the rank of the transpose matrix $A^{\mathrm{T}}$.

### Hint.

Recall that the rank of a matrix $A$ is the dimension of the range of $A$.
The range of $A$ is spanned by the column vectors of the matrix $A$.

### Proof.

We write $A = (a_{ij})$ and let

$$\mathbf{A}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{ni} \end{bmatrix}$$

be the $i$-th column vector of $A$ for $i = 1, 2, \ldots, n$.
Also let

$$\mathbf{B}_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \cdots \\ a_{in} \end{bmatrix}$$

be the $i$-th column vector of $A^{\mathrm{T}}$ for $i = 1, 2, \ldots, m$, that is $B_i$ is the transpose of the $i$-th row vector of $A$ .

---

Suppose that $\operatorname{rank}(A) = \dim(\mathcal{R}(A^{\mathrm{T}})) = k$ and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a basis for the range $\mathcal{R}(A^{\mathrm{T}})$. We write

$$\mathbf{v}_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{in} \end{bmatrix}$$

for $i = 1, 2, \ldots, k$.
Then each column vector $\mathbf{B}_i$ of $A^{\mathrm{T}}$ is a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$. Thus we have

$$\begin{aligned} \mathbf{B}_1 &= c_{11}\mathbf{v}_1 + \cdots + c_{1k}\mathbf{v}_k \\ \mathbf{B}_2 &= c_{21}\mathbf{v}_1 + \cdots + c_{2k}\mathbf{v}_k \\ &\vdots \\ \mathbf{B}_m &= c_{m1}\mathbf{v}_1 + \cdots + c_{mk}\mathbf{v}_k. \end{aligned}$$

---

More explicitly we have,

$$\begin{bmatrix} a_{11} \\ a_{12} \\ \cdots \\ a_{1n} \end{bmatrix} = c_{11} \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix} + \cdots + c_{1k} \begin{bmatrix} v_{k1} \\ v_{k2} \\ \vdots \\ v_{kn} \end{bmatrix}$$

$$\begin{bmatrix} a_{21} \\ a_{22} \\ \cdots \\ a_{2n} \end{bmatrix} = c_{21} \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix} + \cdots + c_{2k} \begin{bmatrix} v_{k1} \\ v_{k2} \\ \vdots \\ v_{kn} \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} a_{m1} \\ a_{m2} \\ \cdots \\ a_{mn} \end{bmatrix} = c_{m1} \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix} + \cdots + c_{mk} \begin{bmatrix} v_{k1} \\ v_{k2} \\ \vdots \\ v_{kn} \end{bmatrix}$$

Now, we look at the $i$-th entries for the above vectors and we have

$$\begin{aligned} a_{1i} &= c_{11}v_{1i} + \cdots + c_{1k}v_{ki} \\ a_{2i} &= c_{21}v_{1i} + \cdots + c_{2k}v_{ki} \\ &\vdots \\ a_{mi} &= c_{m1}v_{1i} + \cdots + c_{mk}v_{ki} \\ &. \end{aligned}$$

We rewrite these as a vector equality and obtain

$$\begin{aligned} \mathbf{A}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{ni} \end{bmatrix} &= v_{1i} \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + \cdots + v_{ki} \begin{bmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{bmatrix} \\ &= v_{1i}\mathbf{c}_1 + \cdots + v_{ki}\mathbf{c}_k, \end{aligned}$$

where we put

$$c_j = \begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{mj} \end{bmatrix}$$

for $j = 1, 2, \ldots, k$.

---

This shows that any column vector $\mathbf{A}_i$ of $A$ is a linear combination of vectors $\mathbf{c}_1, \ldots, \mathbf{c}_k$. Therefore we have

$$\mathcal{R}(A) = \text{Span}\{\mathbf{A}_1, \ldots, \mathbf{A}_n\} \subset \text{Span}\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}.$$

Since the dimension of a subspace is smaller than the dimension of a vector space containing it, we have

$$\text{rank}(A) = \dim(\mathcal{R}(A) \leq \dim(\text{Span}\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}) \leq k.$$

Hence we obtain

$$\text{rank}(A) \leq \text{rank}(A^{\mathrm{T}}).$$

---

To achieve the opposite inequality, we repeat this argument using $A^{\mathrm{T}}$, and we obtain

$$\text{rank}(A^{\mathrm{T}}) \leq \text{rank}((A^{\mathrm{T}})^{\mathrm{T}}) = \text{rank}(A)$$

since we have $(A^{\mathrm{T}})^{\mathrm{T}} = A$.

Therefore, we proved the required equality

$$\operatorname{rank}(A) = \operatorname{rank}(A^{\mathrm{T}}).$$

# Subspaces in General Vector Spaces

## Definition

① If a nonempty subset $W$ of a vector space $V$ is itself a vector space, we call $W$ a **subspace** in $V$.

## Summary

① (**Subspace Criteria**) A subset $W$ of a vector space $V$ is a subspace if and only if

1. The zero vector in $V$ is in $W$.
2. For any vectors $A, B \in W$, the addition $A + B \in W$.
3. For any vector $A \in W$ and a scalar $c$, the scalar multiplication $cA \in W$.

## Problems

① Each of the following sets are not a subspace of the specified vector space. For each set, give a reason why it is not a subspace.

**(1)**

$$S_1 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 \geq 0 \right\}$$

in the vector space $\mathbb{R}^3$.

**(2)**

$$S_2 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 - 4x_2 + 5x_3 = 2 \right\}$$

in the vector space $\mathbb{R}^3$.

**(3)**

$$S_3 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; y = x^2 \right\}$$

in the vector space $\mathbb{R}^2$.

**(4)** Let $P_4$ be the vector space of all polynomials of degree $4$ or less with real coefficients.

$$S_4 = \{f(x) \in P_4 \mid f(1) \text{ is an integer}\}$$

in the vector space $P_4$.

**(5)**

$$S_5 = \{f(x) \in P_4 \mid f(1) \text{ is a rational number}\}$$

in the vector space $P_4$.

**(6)** Let $M_{2\times 2}$ be the vector space of all $2 \times 2$ real matrices.

$$S_6 = \{A \in M_{2\times 2} \mid \det(A) \neq 0\}$$

in the vector space $M_{2\times 2}$.
**(7)**

$$S_7 = \{A \in M_{2\times 2} \mid \det(A) = 0\}$$

in the vector space $M_{2\times 2}$.
**(8)** Let $C[-1,1]$ be the vector space of all real continuous functions defined on the interval $[a,b]$.

$$S_8 = \{f(x) \in C[-2,2] \mid f(-1)f(1) = 0\}$$

in the vector space $C[-2,2]$.
**(9)**

$$S_9 = \{f(x) \in C[-1,1] \mid f(x) \geq 0 \text{ for all } -1 \leq x \leq 1\}$$

in the vector space $C[-1,1]$.
**(10)** Let $C^2[a,b]$ be the vector space of all real-valued functions $f(x)$ defined on $[a,b]$, where $f(x)$, $f'(x)$, and $f''(x)$ are continuous on $[a,b]$. Here $f'(x)$, $f''(x)$ are the first and second derivative of $f(x)$.

$$S_{10} = \{f(x) \in C^2[-1,1] \mid f''(x) + f(x) = \sin(x) \text{ for all } -1 \leq x \leq 1\}$$

in the vector space $C[-1,1]$.
**(11)** Let $S_{11}$ be the set of real polynomials of degree exactly $k$, where $k \geq 1$ is an integer, in the vector space $P_k$.
**(12)** Let $V$ be a vector space and $W \subset V$ a vector subspace. Define the subset $S_{12}$ to be the **complement** of $W$,

$$V \setminus W = \{\mathbf{v} \in V \mid \mathbf{v} \notin W\}.$$

## Solution.

Recall the following subspace criteria.
A subset $W$ of a vector space $V$ over the scalar field $K$ is a subspace of $V$ if and only if the following three criteria are met.

2 The subset $W$ contains the zero vector of $V$.

3 If $u, v \in W$, then $u + v \in W$.

4 If $u \in W$ and $a \in K$, then $au \in W$.

Thus, to prove a subset $W$ is not a subspace, we just need to find a counterexample of any of the three criteria.

### Solution (1). $S_1 = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 \geq 0\}$

The subset $S_1$ does not satisfy condition 3. For example, consider the vector

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Then since $x_1 = 1 \geq 0$, the vector $\mathbf{x} \in S_1$. Then consider the scalar product of $\mathbf{x}$ and the scalar $-1$. Then we have

$$(-1) \cdot \mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix},$$

and the first entry is $-1$, hence $-\mathbf{x}$ is not in $S_1$. Thus $S_1$ does not satisfy condition 3 and it is not a subspace of $\mathbb{R}^3$.

(You can check that conditions 1, 2 are met.)

Solution (2). $S_2 = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 - 4x_2 + 5x_3 = 2\}$

The zero vector of the vector space $\mathbb{R}^3$ is

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Since the zero vector $\mathbf{0}$ does not satisfy the defining relation $x_1 - 4x_2 + 5x_3 = 2$, it is not in $S_2$. Hence condition 1 is not met, hence $S_2$ is not a subspace of $\mathbb{R}^3$.

(You can check that conditions 2, 3 are not met as well.)

Solution (3). $S_3 = \{\mathbf{x} \in \mathbb{R}^2 \mid y = x^2 \quad \}$

Consider vectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

These are vectors in $S_3$ since both vectors satisfy the defining relation $y = x^2$.

However, their sum

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

is not in $S_3$ since $1 \neq 0^2$.

Hence condition 2 is not met, and thus $S_3$ is not a subspace of $\mathbb{R}^2$.

(You can check that condition 1 is fulfilled yet condition 3 is not met.)

Solution (4). $S_4 = \{f(x) \in P_4 \mid f(1) \text{ is an integer}\}$

Consider the polynomial $f(x) = x$. Since the degree of $f(x)$ is $1$ and $f(1) = 1$ is an integer, it is in $S_4$. Consider the scalar product of $f(x)$ and the scalar $1/2 \in \mathbb{R}$.

Then we evaluate the scalar product at $x = 1$ and we have

$$\frac{1}{2}f(1) = \frac{1}{2},$$

which is not an integer.

Thus $(1/2)f(x)$ is not in $S_4$, hence condition 3 is not met. Thus $S_4$ is not a subspace of $P_4$.

(You can check that conditions 1, 2 are met.)

Solution (5). $S_5 = \{f(x) \in P_4 \mid f(1) \text{ is a rational number}\}$

Let $f(x) = x$. Then $f(x)$ is a degree $1$ polynomial and $f(1) = 1$ is a rational number.

However, the scalar product $\sqrt{2}f(x)$ of $f(x)$ and the scalar $\sqrt{2} \in \mathbb{R}$ is not in $S_5$ since

$$\sqrt{2}f(1) = \sqrt{2},$$

which is not a rational number. Hence condition 3 is not met and $S_5$ is not a subspace of $P_4$.

(You can check that conditions 1, 2 are met.)

Solution (6). $S_6 = \{A \in M_{2\times 2} \mid \det(A) \neq 0\}$

The zero vector of the vector space $M_{2\times 2}$ is the $2 \times 2$ zero matrix $O$.

Since the determinant of the zero matrix $O$ is $0$, it is not in $S_6$. Thus, condition 1 is not met and $S_6$ is not a subspace of $M_{2\times 2}$.

(You can check that conditions 2, 3 are not met as well.)

Solution (7). $S_7 = \{A \in M_{2\times 2} \mid \det(A) = 0\}$

Consider the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The determinants of $A$ and $B$ are both $0$, hence they belong to $S_7$.

However, their sum

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

has the determinant $1$, hence the sum $A + B$ is not in $S_7$.

So condition 2 is not met and $S_7$ is not a subspace of $M_{2\times 2}$.

(You can check that conditions 1, 3 are met.)

Solution (8). $S_8 = \{f(x) \in C[-2, 2] \mid f(-1)f(1) = 0\}$

Consider the continuous functions

$$f(x) = x - 1 \text{ and } g(x) = x + 1.$$

(These are polynomials, hence they are continuous.)

We have

$$\begin{aligned} f(-1)f(1) &= (-2)\cdot(0) = 0 \text{ and} \\ g(-1)g(1) &= (0)\cdot 2 = 0. \end{aligned}$$

So these functions are in $S_8$.

However, their sum $h(x) := f(x) + g(x)$ does not belong to $S_8$ since we have

$$\begin{aligned} h(-1)h(1) &= \big(f(-1)+g(-1)\big)\big(f(1)+g(1)\big) \\ &= (-2+0)(0+2) = -4 \neq 0. \end{aligned}$$

Therefore, condition 2 is not met and $S_8$ is not a subspace of $C[-1,1]$.

(You can check that conditions 1, 3 are met.)

> Solution (9). $S_9 = \{f(x) \in C[-1,1] \mid f(x) \geq 0 \text{ for all } -1 \leq x \leq 1\}$

Let $f(x) = x^2$, an open-up parabola.

Then $f(x)$ is continuous and non-negative for $-1 \leq x \leq 1$. Hence $f(x) = x^2$ is in $S_9$.

However, the scalar product $(-1)f(x)$ of $f(x)$ and the scalar $-1$ is not in $S_9$ since, say,

$$(-1)f(1) = -1$$

is negative.

So condition 3 is not met and $S_9$ is not a subspace of $C[-1,1]$.

(You can check that conditions 1, 2 are met.)

> Solution (10).
> $S_{10} = \{f(x) \in C^2[-1,1] \mid f''(x) + f(x) = \sin(x) \text{ for all } -1 \leq x \leq 1\}$

The zero vector of the vector space $C^2[-1,1]$ is the zero function $\theta(x) = 0$.

The second derivative of the zero function is still the zero function.

Thus,

$$\theta''(x) + \theta(x) = 0$$

and since $\sin(x)$ is not the zero function, $\theta(x)$ is not in $S_{10}$.

Hence $S_{10}$ is not a subspace of $C^2[-1,1]$.

(You can check that conditions 2, 3 are not met as well.

For example, consider the function $f(x) = -\frac{1}{2}x\cos(x) \in S_{10}$.)

> Solution (11). Let $S_{11}$ be the set of real polynomials of degree exactly $k$.

The set $S_{11}$ is not a vector subspace of $\mathbf{P}_k$. One reason is that the zero function $\mathbf{0}$ has degree $0$, and so does not lie in $S_{11}$. The set $S_{11}$ is also not closed under addition. Consider the two polynomials

$f(x) = x^k + 1$ and $g(x) =-x^k + 1$. Both of these polynomials lie in $S_{11}$, however $f(x) + g(x) = 2$ has degree $0$ and so does not lie in $S_{11}$.

### Solution (12). The complement

The complement $S_{12} = V \setminus W$ is not a vector subspace. Specifically, if $\mathbf{0} \in V$ is the zero vector, then we know $\mathbf{0} \in W$ because $W$ is a subspace. But then $\mathbf{0} \notin V \setminus W$, and so $V \setminus W$ cannot be a vector subspace.

---

⑤ Let $V$ be the vector space over $\mathbb{R}$ of all real valued functions defined on the interval $[0, 1]$. Determine whether the following subsets of $V$ are subspaces or not.
**(a)** $S = \{f(x) \in V \mid f(0) = f(1)\}$.
**(b)** $T = \{f(x) \in V \mid f(0) = f(1) + 3\}$.

---

## Hint.

To show that a subset $W$ of a vector space $V$ is a subspace, we need to check that

⑥ the zero vector in $V$ is in $W$

⑦ for any two vectors $u, v \in W$, we have $u + v \in W$

⑧ for any scalar $c$ and any vector $u \in W$, we have $cu \in W$.

## Solution.

### (a) Is $S = \{f(x) \in V \mid f(0) = f(1)\}$ a subspace?

We show that $S$ is a subspace of the vector space $V$ by checking conditions (1)-(3) given in the hint above.
First note that the zero vector in $V$ is the zero function $\theta(x)$, that is, $\theta(x) = 0$ for any $x \in [0, 1]$.
Since we have $\theta(0) = 0 = \theta(1)$, the zero function $\theta(x) \in S$.
Condition (1) is met.

Now, take any $f(x), g(x) \in S$. By the defining relation of $S$, we have

$$f(0) = f(1), \quad g(0) = g(1).$$

Consider the addition $(f + g)(x)$. We have

$$(f + g)(0) = f(0) + g(0) = f(1) + g(1) = (f + g)(1)$$

and it follows that $(f + g)(x) \in S$.
Thus $S$ satisfies condition (2).

To check the condition (3), take any scalar $c \in \mathbb{R}$ and $f(x) \in S$ .
Since $f(x) \in S$ , we have $f(0) = f(1)$ . The scalar multiplication $(cf)(x)$ satisfies

$$(cf)(0) = c \cdot f(0) = c \cdot f(1) = (cf)(0).$$

Thus $(cf)(x) \in S$ .

Therefore, the subset $S$ satisfies conditions (1)-(3). Hence $S$ is a subspace of $V$.

### (b) Is $T = \{f(x) \in V \mid f(0) = f(1) + 3\}$ a subspace?

We claim that $T$ is not a subspace of the vector space $V$.
For example, the subset $T$ does not satisfy condition (1).

The zero vector of $V$ is the zero function $\theta(x)$.
Then we have

$$\theta(0) = 0 \neq 0 + 3 = \theta(1) + 3,$$

and hence the zero vector $\theta(x) \in V$ is not in $W$.

⑨ Let $V$ be the vector space over $\mathbb{R}$ of all real valued function on the interval $[0, 1]$ and let

$$W = \{f(x) \in V \mid f(x) = f(1 - x) \text{ for } x \in [0, 1]\}$$

be a subset of $V$. Determine whether the subset $W$ is a subspace of the vector space $V$.

### Proof.

We claim that $W$ is a subspace of $V$.
To show the claim, we need to check that the following subspace criteria.

**Subspace Criteria.**

⑩ The zero vector in $V$ is in $W$.

⑪ For any two elements $f(x), g(x) \in W$ , we have $f(x) + g(x) \in W$ .

⑫ For any scalar $c$ and any element $f(x) \in W$ , we have $cf(x) \in W$ .

The zero vector of $V$ is the zero function $\theta(x) = 0$.
Since we have

$$\theta(x) = 0 = \theta(1-x)$$

for any $x \in [0,1]$ , the zero vector $\theta$ is in $W$ , hence condition 1 is met.

---

Let $f(x), g(x)$ be arbitrary elements in $W$ . Then these functions satisfy

$$f(x) = f(1-x) \text{ and } g(x) = g(1-x)$$

for any $x \in [0,1]$ .
We want to show that the sum $h(x) := f(x) + g(x)$ is in $W$ . This follows since we have

$$h(x) = f(x) + g(x) = f(1-x) + g(1-x) = h(1-x).$$

Thus, condition 2 is satisfied.

---

Finally, we check condition 3. Let $c$ be a scalar and let $f(x)$ be an element in $W$ .
Then we have

$$f(x) = f(1-x).$$

It follows from this that

$$cf(x) = cf(1-x),$$

and this shows that the scalar product $cf(x)$ is in $W$ .

Therefore condition 3 holds, and we have proved the subspace criteria for $W$ . Thus $W$ is a subspace of the vector space $V$ .

⑬ Let $C[-1,1]$ be the vector space over $\mathbb{R}$ of all real-valued continuous functions defined on the interval $[-1,1]$. Consider the subset $F$ of $C[-1,1]$ defined by

$$F = \{f(x) \in C[-1,1] \mid f(0) \text{ is an integer}\}.$$

Prove or disprove that $F$ is a subspace of $C[-1,1]$.

## Proof.

### (a) Continuous functions taking integer values at $0$.

We claim that the subset $F$ is not a subspace of the vector space $C[-1,1]$ .
If $F$ is a subspace, then $F$ is closed under scalar multiplication.
For instance, consider the scalar multiplication of the function $f(x) = x+1$ and the scalar $r = 1/2$ .
Note that since $f(0) = 1$ is an integer, the function is in $F$ .
However, the scalar multiplication

$$rf(x) = \frac{1}{2}(x+1)$$

is not in $F$ because

$$rf(0) = \frac{1}{2}$$

is not an integer.

Therefore, $F$ is not closed under scalar multiplication, hence $F$ is not a subspace of $C[-1,1]$.

### (b) Set of all $n \times n$ skew-symmetric matrices

We prove that $W$ is a subspace of $V$ by showing the following subspace criteria.

⑭ The zero vector $0 \in M_{n\times n}$ is in $W$.

⑮ If $A, B \in W$, then $A + B \in W$.

⑯ If $A \in W, r \in \mathbb{R}$, then $rA \in W$.

First, note that the zero vector in the vector space $M_{n\times n}$ is the $n \times n$ zero matrix $O$. Since we have

$$O^{\mathrm{T}} = O = -O,$$

the zero matrix $O$ is skew-symmetric, hence the zero vector $O$ is in $W$. So condition 1 is met.

To prove condition 2, let $A, B$ be arbitrary elements in $W$. Thus, $A$ and $B$ are skew-symmetric matrices:

$$A^{\mathrm{T}} = -A \text{ and } B^{\mathrm{T}} = -B. \tag{*}$$

We want to prove that the addition $A + B$ is in $W$, namely, we want to show $A + B$ is skew-symmetric.

We have

$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}} \overset{(*)}{=} -A + (-B) = -(A+B),$$

and this implies that $A + B$ is skew-symmetric. Hence $A + B \in W$ and condition 2 is satisfied.

Finally, we check condition 3. Let $A \in W$ and $r \in \mathbb{R}$. We want to show that the scalar product $rA \in W$, that is, $rA$ is skew-symmetric.

We have

$$\begin{aligned}(rA)^{\mathrm{T}} &= rA^{\mathrm{T}} \\ &= r(-A) && (A \text{ is skew-symmetric}) \\ &= -rA,\end{aligned}$$

which yields that $rA$ is skew-symmetric. Hence $rA \in W$ and condition 3 is met as well.

Thus, we have checked all the subspace criteria. Hence $W$ is a subspace of $M_{n\times n}$.

⑰ Let $P_3$ be the vector space over $\mathbb{R}$ of all degree three or less polynomial with real number coefficient. Let $W$ be the following subset of $P_3$.

$$W = \{p(x) \in P_3 \mid p'(-1) = 0 \text{ and } p''(1) = 0\}.$$

Here $p'(x)$ is the first derivative of $p(x)$ and $p''(x)$ is the second derivative of $p(x)$. Show that $W$ is a subspace of $P_3$ and find a basis for $W$.

## Proof.

### Subspace criteria

To show that the subset $W$ of the vector space $P_3$ is a subspace, we need to check that

⑱ the zero vector in $P_3$ is in $W$

⑲ for any two vectors $u, v \in W$, we have $u + v \in W$

⑳ for any scalar $c$ and any vector $u \in W$, we have $cu \in W$.

### Check Condition 1

First note that the zero vector in $P_3$ is the zero polynomial, which we denote $\theta(x)$.
Thus we have $\theta(x) = 0$ for any $x$.

The derivative of the zero function is still the zero function, we have

$$\theta'(x) = 0, \text{ and } \theta''(x) = 0.$$

Therefore the zero vector $\theta(x)$ satisfies the following defining relations for $W$

$$\theta'(-1) = 0 \text{ and } \theta''(1) = 0$$

and thus the zero vector $\theta(x)$ is in $W$.
Thus condition 1 is met.

### Check Condition 2

To check condition 2, let $f(x), g(x) \in W$.
Then $f(x)$ and $g(x)$ satisfy

$$f'(-1) = 0, f''(1) = 0, \text{ and } g'(-1) = 0, g''(1) = 0.$$

Let $h(x) := f(x) + g(x)$. We want to show that the sum $h(x) = f(x) + g(x)$ is in $W$.

We have

$$h'(x) = (f(x) + g(x))' = f'(x) + g'(x)$$

and similarly we have

$$h''(x) = f''(x) + g''(x).$$

Thus we obtain

$$h'(-1) = f'(-1) + g'(-1) = 0 + 0 = 0 \text{ and } h''(1) = f''(1) + g''(1) = 0 + 0 = 0.$$

Therefore $h(x)$ satisfies the defining relations of $W$ and hence $h(x) = f(x) + g(x) \in W$ .
So condition 2 is satisfied.

### Check Condition 3

Condition 3 can be checked as follows. Let $f(x) \in W$ and let $r \in R$ .
We want to show that the scalar product $k(x) := rf(x)$ is in $W$ .

Since $f(x)$ is in $W$ , we have

$$f'(-1) = 0, f''(1) = 0.$$

Note that we have

$$k'(x) = (rf(x))' = rf'(x), \text{ and } k''(x) = rf''(x).$$

Thus we see that

$$k'(-1) = rf'(-1) = r \cdot 0 = 0 \text{ and } k''(1) = rf''(1) = r \cdot 0 = 0.$$

Thus $k(x)$ satisfies the defining relations of $W$ and hence the scalar product $k(x) = rf(x)$ is in $W$ .
Therefore condition 3 and thus all conditions are met, and we conclude that $W$ is a subspace of $P_3$ .

### Find a basis for the subspace $W$ .

To find a basis, we observe the following.

Any vector in $W$ is a polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$$

satisfying

$$p'(-1) = 0 \text{ and } p''(1) = 0.$$

Since we have

$$p'(x) = a_1 + 2a_2x + 3a_3x^2 \text{ and } p''(x) = 2a_2 + 6a_3x,$$

the above conditions become

$$a_1 - 2a_2 + 3a_3 = 0 \text{ and } 2a_2 + 6a_3 = 0.$$

To find solutions for these equations consider the augmented matrix

$$\left[\begin{array}{rrr|r} 1 & -2 & 3 & 0 \\ 0 & 2 & 6 & 0 \end{array}\right].$$

We apply the elementary row operations to this matrix and obtain the following reduced row echelon matrix.

$$\left[\begin{array}{ccc|c} 1 & 0 & 9 & 0 \\ 0 & 1 & 3 & 0 \end{array}\right].$$

Thus, solutions are

$$\begin{aligned} a_1 &= -9a_3 \\ a_2 &= -3a_3. \end{aligned}$$

---

Therefore, any polynomial $p(x)$ in $W$ can be written as

$$\begin{aligned} p(x) &= a_0 - 9a_3x - 3a_3x^2 + a_3x^3 \\ &= a_0(1) + a_3(-9x - 3x^2 + x^3). \end{aligned}$$

In particular we see that the polynomials $q_1(x) := 1$ and $q_2(x) := -9x - 3x^2 + x^3$ are in $W$. Since any vector $p(x) \in W$ is a linear combination of $q_1(x)$ and $q_2(x)$, the set $\{q_1(x), q_2(x)\}$ is a spanning set for $W$.

---

We check that $q_1(x)$ and $q_2(x)$ are linearly independent.
If we have a linear combination

$$c_1q_1(x) + c_2q_2(x) = 0(= \theta(x)),$$

then we have

$$c_1 - 9c_2x - 3c_2x^2 + c_2x^3 = 0.$$

Thus we see that $c_1 = c_2 = 0$ and the vectors $q_1(x)$ and $q_2(x)$ are linearly independent.
Therefore $\{q_1(x), q_2(x)\}$ is a linearly independent spanning set for $W$, and thus it is a basis for $W$.

---

In summary we found a basis

$$\{1, -9x - 3x^2 + x^3\}.$$

Remark: The dimension of $W$ is $2$.

(21) Let $V$ be the vector space over $\mathbb{R}$ consisting of all $n \times n$ real matrices for some fixed integer $n$. Prove or disprove that the following subsets of $V$ are subspaces of $V$.
**(a)** The set $S$ consisting of all $n \times n$ symmetric matrices.
**(b)** The set $T$ consisting of all $n \times n$ skew-symmetric matrices.
**(c)** The set $U$ consisting of all $n \times n$ nonsingular matrices.

### Hint.

Recall that

(22) a matrix $A$ is symmetric if $A^{\mathrm{T}} = A$.

㉓ a matrix $A$ is skew-symmetric if $A^{\mathrm{T}} = -A$.

## Proof.

To show that a subset $W$ of a vector space $V$ is a subspace, we need to check that

㉔ the zero vector in $V$ is in $W$

㉕ for any two vectors $u, v \in W$, we have $u + v \in W$

㉖ for any scalar $c$ and any vector $u \in W$, we have $cu \in W$.

### (a) The set $S$ consisting of all $n \times n$ symmetric matrices.

We will prove that $S$ is a subspace of $V$. The zero vector $O$ in $V$ is the $n \times n$ zero matrix and it is symmetric. Thus the zero vector $O \in S$ and the condition 1 is met.

To check the second condition, take any $A, B \in S$, that is, $A, B$ are symmetric matrices.
To show that $A + B \in S$, we need to check that the matrix $A + B$ is symmetric.

We have

$$(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}} = A + B$$

since $A, B$ are symmetric. Thus $A + B$ is also symmetric, and $A + B \in S$. Condition 2 is also satisfied.

Finally, to check condition 3, let $A \in S$ and let $r \in R$. We show that $rA \in S$, namely, we show that $rA$ is symmetric.
We have

$$(rA)^{\mathrm{T}} = rA^{\mathrm{T}} = rA$$

since $A$ is symmetric. Thus $rA$ is symmetric and hence $rA \in S$.
Thus condition 3 is met.
By the subspace criteria, the subset $S$ is a subspace of the vector space $V$.

### (b) The set $T$ consisting of all $n \times n$ skew-symmetric matrices.

We will prove that $T$ is a subspace of $V$.
The zero vector $O$ in $V$ is the $n \times n$ matrix, and it is skew-symmetric because

$$O^{\mathrm{T}} = O = -O.$$

Thus condition 1 is met.

---

For condition 2, take arbitrary elements $A, B \in T$. The matrices $A, B$ are skew-symmetric, namely, we have

$$A^{\mathrm{T}} = -A \text{ and } B^{\mathrm{T}} = -B.$$

We show that $A + B \in T$, or equivalently we show that the matrix $A + B$ is skew-symmetric.

We have

$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}} \overset{*}{=} -A + (-B) = -(A+B).$$

Therefore the matrix $A + B$ is skew-symmetric and condition 2 is met.

---

To prove the last condition, consider any $A \in T$ and $r \in \mathrm{R}$.
We show that $rA$ is skew-symmetric, and hence $rA \in T$.
Using the fact that $A$ is skew-symmetric ($A^{\mathrm{T}} = -A$), we have

$$(rA)^{\mathrm{T}} = rA^{\mathrm{T}} = r(-A) = -rA.$$

Hence $rA$ is skew-symmetric and condition 3 is satisfied.

By the subspace criteria, the subset $T$ is a subspace of the vector space $V$.

> (c) The set $U$ consisting of all $n \times n$ nonsingular matrices.

We claim that $U$ is not a subspace of $V$.
As the zero vector of $V$ is the $n \times n$ matrix and the zero matrix is singular, the zero vector is not in $U$.
Hence condition 1 is not met, and thus $U$ is not a subspace.

---

Another reason that $U$ is not a subspace is that the addition is not closed. For example,
if $A$ is a nonsingular matrix (say, $A$ is $n \times n$ identity matrix), then $-A$ is also nonsingular matrix but their addition $A + (-A) = O$ is nonsingular, hence it is not in $U$.

---

**27** Let $V$ be the vector space of $n \times n$ matrices with real coefficients, and define

$$W = \{\mathbf{v} \in V \mid \mathbf{v}\mathbf{w} = \mathbf{w}\mathbf{v} \text{ for all } \mathbf{w} \in V\}.$$

The set $W$ is called the **center** of $V$. Prove that $W$ is a subspace of $V$.

**Proof.**

We must show that $W$ satisfies the three criteria for vector subspaces.
Namely, the zero vector of $V$ is in $W$, and $W$ is closed under addition and scalar multiplication.

First, the zero element in $V$ is the matrix $0$ whose entries are all $0$. For any other matrix $\mathbf{x} \in V$, we have $0\mathbf{x} = 0 = \mathbf{x}0$. So we see that $0 \in W$.

---

Now suppose $\mathbf{v}, \mathbf{w} \in W$ and $c \in \mathbb{R}$. Then for any $\mathbf{x} \in V$, we have

$$(\mathbf{v} + \mathbf{w})\mathbf{x} = \mathbf{v}\mathbf{x} + \mathbf{w}\mathbf{x} = \mathbf{x}\mathbf{v} + \mathbf{x}\mathbf{w} = \mathbf{x}(\mathbf{v} + \mathbf{w}),$$

where the second equality follows because $\mathbf{v}$ and $\mathbf{w}$ lie in $W$. So we see that $\mathbf{v} + \mathbf{w} \in W$ as well, and so $W$ is closed under addition.

---

Finally we must show that $c\mathbf{v} \in W$ as well. For any other $\mathbf{x} \in V$, we have

$$(c\mathbf{v})\mathbf{x} = c(\mathbf{v}\mathbf{x}) = c(\mathbf{x}\mathbf{v}) = \mathbf{x}(c\mathbf{v}),$$

where the second equality follows from the fact that $\mathbf{v} \in W$ and so $\mathbf{v}\mathbf{x} = \mathbf{x}\mathbf{v}$.

Thus we see that $c\mathbf{v} \in W$, finishing the proof.

㉘ Let $C(\mathbb{R})$ be the vector space of real-valued functions on $\mathbb{R}$. Consider the set of functions $W = \{f(x) = a + b\cos(x) + c\cos(2x) \mid a, b, c \in \mathbb{R}\}$. Prove that $W$ is a vector subspace of $C(\mathbb{R})$.

**Proof.**

We verify the subspace criteria: the zero vector of $C(\mathbb{R})$ is in $W$, and $W$ is closed under addition and scalar multiplication.

---

First, the zero element of $C(\mathbb{R})$ is the zero function $\mathbf{0}$ defined by $\mathbf{0}(x) = 0$. This element lies in $W$, as $\mathbf{0}(x) = 0 + 0\cos(x) + 0\cos(2x)$.

---

Now suppose $f_1(x), f_2(x) \in W$, say $f_1(x) = a_1 + b_1\cos(x) + c_1\cos(2x)$ and $f_2(x) = a_2 + b_2\cos(x) + c_2\cos(2x)$. Then

$$f_1(x) + f_2(x) = (a_1 + a_2) + (b_1 + b_2)\cos(x) + (c_1 + c_2)\cos(2x)$$

and so $f_1(x) + f_2(x) \in W$. Finally, for any scalar $d \in \mathbb{R}$, we have

$$df_1(x) = (a_1 d) + (b_1 d)\cos(x) + (c_1 d)\cos(2x),$$

and so $df_1(x) \in W$ as well.

This proves that $W$ is a subspace of $W$.

29 For what real values of $a$ is the set

$$W_a = \{f \in C(\mathbb{R}) \mid f(0) = a\}$$

a subspace of the vector space $C(\mathbb{R})$ of all real-valued functions?

## Solution.

The zero element of $C(\mathbb{R})$ is the function $0$ defined by $0(x) = 0$. This shows that $0 \in W_a$ if and only if $a = 0$.

We have shown that if $a \neq 0$, then $W_a$ is not a subspace as every subspace contains the zero vector. Now we consider the case $a = 0$ and prove that $W_0$ is a subspace.

We verify the subspace criteria: the zero vector of $C(\mathbb{R})$ is in $W_0$, and $W_0$ is closed under addition and scalar multiplication.

As mentioned before, $0 \in W_0$.

Now suppose $f, g \in W_0$. Then $f(0) = g(0) = 0$, and so

$$(f+g)(0) = f(0) + g(0) = 0.$$

Thus $f + g \in W_0$. Finally, if $c \in \mathbb{R}$ is a scalar and $f \in W_0$, then

$$(cf)(0) = cf(0) = c \cdot 0 = 0.$$

Thus $cf \in W_0$, and $W_0$ is a vector subspace.

30 Let $V$ be the vector space of $n \times n$ matrices, and $M \in V$ a fixed matrix. Define

$$W = \{A \in V \mid AM = MA\}.$$

The set $W$ here is called the **centralizer** of $M$ in $V$. Prove that $W$ is a subspace of $V$.

## Proof.

First we check that the zero element of $V$ lies in $W$. The zero element of $V$ is the $n \times n$ zero matrix $0$.

It is clear that $M0 = 0 = 0M$, and so $0 \in W$.

Next suppose $A, B \in W$ and $c \in \mathbb{R}$. Then $AM = MA$ and $BM = MB$, and so

$$(A+B)M = AM + BM = MA + MB = M(A+B).$$

Thus, $A + B \in W$.

We also have

$$(cA)M = c(AM) = c(MA) = M(cA),$$

and so $cA \in W$ .

These three criteria show that $W$ is a subspace of $V$.

31 Let $V$ be the vector space of $k \times k$ matrices. Then for fixed matrices $R, S \in V$, define the subset $W = \{RAS \mid A \in V\}$. Prove that $W$ is a vector subspace of $V$.

### Proof.

We verify the subspace criteria: the zero vector of $V$ is in $W$, and $W$ is closed under addition and scalar multiplication.

First, let $\mathbf{0} \in V$ be the $k \times k$ zero matrix. Then $R\mathbf{0}S = \mathbf{0}$, and so $\mathbf{0} \in W$.

Now suppose $X, Y \in W$. Then there are elements $A, B \in V$ such that $RAS = X$ and $RBS = Y$. Then

$$X + Y = RAS + RBS = R(A + B)S$$

and so $X + Y \in W$.

Now for a scalar $c \in \mathbb{R}$ and matrix $X = RAS \in W$, we have

$$cX = cRAS = R(cA)S,$$

and so $cX \in W$ as well.

32 Let $V$ denote the vector space of all real $n \times n$ matrices, where $n$ is a positive integer. Determine whether the set $U$ of all $n \times n$ nilpotent matrices is a subspace of the vector space $V$ or not.

### Solution.

We claim that the matrix $B - A$ is not necessarily invertible.

Consider the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

This matrix is nilpotent as we have

$$A^2 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Also consider the matrix

$$B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Since the determinant of the matrix $B$ is $1$, it is invertible.

So the matrix $A$ and $B$ satisfy the assumption of the problem.

However the matrix

$$B - A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

is not invertible as its determinant is $0$.

Hence we found a counterexample.

## Related Question.

Here is another problem about a nilpotent matrix.

**Problem.**

Let $A$ be an $n \times n$ nilpotent matrix. Then prove that $I - A, I + A$ are both nonsingular matrices, where $I$ is the $n \times n$ identity matrix.

㉝ Let $V$ be a vector space over a scalar field $K$.
Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ be vectors in $V$ and consider the subset

$$W = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k \mid a_1, a_2, \dots, a_k \in K \text{ and } a_1 + a_2 + \cdots + a_k = 0\}.$$

So each element of $W$ is a linear combination of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ such that the sum of the coefficients is zero.
Prove that $W$ is a subspace of $V$.

## Proof 1. (Subspace Criteria)

We use the following subspace criteria.

The subset $W$ is a subspace of $V$ if the following three conditions are met.

㉞ The zero vector in $V$ is in $W$.

㉟ For any two elements $\mathbf{v}, \mathbf{v}' \in W$, we have $\mathbf{v} + \mathbf{v}' \in W$.

㊱ For any scalar $c \in K$ and any element $\mathbf{v} \in W$, we have $c\mathbf{v} \in W$.

The zero vector $\mathbf{0}$ of $V$ can be written as

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k.$$

Clearly the sum of the coefficient is zero, hence $\mathbf{0} \in W$.

So condition 1 is met.

---

To verify condition 2, let

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k$$

and

$$\mathbf{v}' = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k$$

be arbitrary elements in $W$. Thus

$$a_1 + a_2 + \cdots + a_k = 0 \text{ and } b_1 + b_2 + \cdots + b_k = 0. \tag{*}$$

The sum $\mathbf{v} + \mathbf{v}'$ is

$$\begin{aligned}\mathbf{v} + \mathbf{v}' &= (a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k) + (b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k)\\ &= (a_1 + b_1)\mathbf{v}_1 + (a_2 + b_2)\mathbf{v}_2 + \cdots + (a_k + b_k)\mathbf{v}_k.\end{aligned}$$

The the sum of the coefficients of the above linear combination is

$$\begin{aligned}&(a_1 + b_1) + (a_2 + b_2) + \cdots + (a_k + b_k)\\ &= (a_1 + a_2 + \cdots + a_k) + (b_1 + b_2 + \cdots + b_k) \overset{(*)}{=} 0 + 0 = 0.\end{aligned}$$

It follows that the sum $\mathbf{v} + \mathbf{v}'$ is in $W$, and hence condition 2 is met.

---

Finally, let us check condition 3. Let

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k$$

be an arbitrary vector in $W$ and let $c \in K$.

Since $\mathbf{v} \in W$, we have

$$a_1 + a_2 + \cdots + a_k = 0.$$

Then the scalar product is

$$\begin{aligned}c\mathbf{v} &= c(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k)\\ &= ca_1\mathbf{v}_1 + ca_2\mathbf{v}_2 + \cdots + ca_k\mathbf{v}_k.\end{aligned}$$

The sum of the coefficients of the linear combination is

$$ca_1 + ca_2 + \cdots + ca_k = c(a_1 + a_2 + \cdots + a_k) = a0 = 0.$$

Hence $c\mathbf{v} \in V$, and condition 3 is met.

Therefore by the subspace criteria, we conclude that $W$ is a subspace of $V$.

### Proof 2. (Span)

Consider an arbitrary vector in $W$:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k \text{ with } a_1 + a_2 + \cdots + a_k = 0.$$

Substituting the relation $a_k = -(a_1 + a_2 + \cdots + a_{k-1})$, we obtain

$$\begin{aligned}
&a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_k\mathbf{v}_k \\
&\quad = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{k-1}\mathbf{v}_{k-1} - (a_1 + a_2 + \cdots + a_{k-1})\mathbf{v}_k \\
&\quad = a_1(\mathbf{v}_1 - \mathbf{v}_k) + a_2(\mathbf{v}_2 - \mathbf{v}_k) + \cdots + a_{k-1}(\mathbf{v}_{k-1} - \mathbf{v}_{k-1}).
\end{aligned}$$

This computation yields that every vector in $W$ is a linear combination of vectors in

$$S := \{\mathbf{v}_1 - \mathbf{v}_k, \mathbf{v}_2 - \mathbf{v}_k, \dots, \mathbf{v}_{k-1} - \mathbf{v}_{k-1}\}.$$

That is, we have $W \subset \text{Span}(S)$.

---

On the other hand, let

$$\mathbf{v} = c_1(\mathbf{v}_1 - \mathbf{v}_k) + c_2(\mathbf{v}_2 - \mathbf{v}_k) + \cdots + c_{k-1}(\mathbf{v}_{k-1} - \mathbf{v}_{k-1})$$

be an arbitrary vector in $\text{Span}(S)$.
Then we have

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{k-1}\mathbf{v}_{k-1} - (c_1 + c_2 + \cdots + c_{k-1})\mathbf{v}_k.$$

This is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$, and the sum of the coefficients is

$$c_1 + c_2 + \cdots + c_{k-1} - (c_1 + c_2 + \cdots + c_{k-1}) = 0.$$

Therefore $\mathbf{v} \in W$. Thus we also have $\text{Span}(S) \subset W$.

---

Putting together these inclusion yields that $W = \text{Span}(S)$.
As the span is always a subspace, we conclude that $W$ is a subspace of $V$.

㊲ Suppose that $S$ is a fixed invertible $3$ by $3$ matrix. This question is about all the matrices $A$ that are diagonalized by $S$ so that $S^{-1}AS$ is diagonal. Show that these matrices $A$ form a subspace of $3$ by $3$ matrix space.
(*MIT*)

### Hint.

Check the following criteria for a subset to be a subspace.

**Theorem. (Subspace criteria)**

A subset $W$ of a vector space $V$ is a subspace if and only if

㊳ The zero vector in $V$ is in $W$.

㊴ For any vectors $A, B \in W$, the addition $A + B \in W$.

㊵ For any vector $A \in W$ and a scalar $c$, the scalar multiplication $cA \in W$.

## Proof.

Let $V$ be the vector space of all $3$ by $3$ matrices.
Define the subset of $V$

$$W = \{A \in V \mid S^{-1}AS \text{ is diagonal.}\}.$$

We want to show that $W$ is a subspace of $V$.

To do this, it suffices to show the following subspace criteria;

㊶ The zero vector in $V$ is in $W$.

㊷ For any vectors $A, B \in W$, the addition $A + B \in W$.

㊸ For any vector $A \in W$ and a scalar $c$, the scalar multiplication $cA \in W$.

The zero vector in the vector space $V$ is the $3$ by $3$ zero matrix $O$.
Since $S^{-1}OS = O$, we have $O \in W$.

To show the second criterion, take $A, B \in W$.
Then we have $S^{-1}(A+B)S = S^{-1}AS + S^{-1}BS$ and since $A, B \in W$, the matrices $S^{-1}AS, S^{-1}BS$ are diagonal.

The addition of diagonal matrices is still diagonal, thus $S^{-1}(A+B)S$ is diagonal and $A + B \in W$.

We finally check the third criterion.

Take $A \in W$ and let $c$ be a scalar.

Then we have $S^{-1}(cA)S = cS^{-1}AS$ . The matrix $S^{-1}AS$ is diagonal since $A \in W$ .

Since a scalar multiplication of a diagonal matrix is still diagonal, we conclude that the matrix $S^{-1}(cA)S$ is diagonal and $cA \in W$ .

Therefore the subset $W$ of $V$ satisfies the criteria to be a subspace.

(44) Let $\mathbf{P}_2$ be the vector space of polynomials of degree $2$ or less. Prove that the set $\{1, 1+x, (1+x)^2\}$ is a basis for $\mathbf{P}_2$.

## Proof.

### (a) Prove that the set $\{1, 1+x, (1+x)^2\}$ is a basis for $\mathbf{P}_2$.

Consider the standard basis $\mathrm{B} = \{1, x, x^2\}$ of $\mathbf{P}_2$. Using this basis, we can write the elements using coordinate vectors as

$$[1]_{\mathrm{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad [1+x]_{\mathrm{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad [(1+x)^2]_{\mathrm{B}} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

We find the coordinate vector by writing an element as a linear combination of the basis elements. For example, $(1+x)^2 = 1 + 2x + 1x^2$, and so the coefficients $1, 2, 1$ translate into the column vector $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

---

Because $\dim \mathbf{P}_2 = 3$, this set is a basis if and only if these three vectors are linearly independent. To verify this, consider the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

This matrix is upper-triangular, and the diagonal entries are all non-zero. This implies the matrix is non-singular, and so the columns are linearly independent.

Thus, the set $\{1, 1+x, (1+x)^2\}$ is a basis of $\mathbf{P}_2$.

### (b) Write the polynomial $f(x) = 2 + 3x - x^2$ as a linear combination of the basis $\{1, 1+x, (1+x)^2\}$.

First Method: Using the technique of completing the square, we can factor the polynomial $f(x)$ as we like. Specifically,

$$\begin{aligned} f(x) &= -x^2 + 3x + 2 \\ &= -(x+1)^2 + 5x + 3 \\ &= -(x+1)^2 + 5(x+1) - 2. \end{aligned}$$

Hence, we have the linear combination

$$f(x) = -2 \cdot 1 + 5(1+x) - (1+x)^2. \tag{*}$$

---

Second Method: We can find this factorization by calculating the Taylor polynomial of $f(x)$ centered at $-1$.

This Taylor polynomial is defined by

$$f(x) = f(-1) + f'(-1)(x+1) + \frac{f''(-1)}{2}(x+1)^2.$$

The polynomial in (*) is recovered by finding $f(-1) = -2$, $f'(-1) = 5$, and $f''(-1) = -2$.

㊺ Let $V$ be a real vector space of all real sequences

$$(a_i)_{i=1}^{\infty} = (a_1, a_2, \cdots).$$

Let $U$ be the subset of $V$ defined by

$$U = \{(a_i)_{i=1}^{\infty} \in V \mid a_{k+2} - 5a_{k+1} + 3a_k = 0, k = 1, 2, \ldots\}.$$

Prove that $U$ is a subspace of $V$.

## Proof.

We prove the subspace criteria.

**Subspace criteria**

㊻ The zero vector in $V$ is in $U$.

㊼ For any two elements $(a_i)_{i=1}^{\infty}, (b_i)_{i=1}^{\infty} \in U$, we have $(a_i)_{i=1}^{\infty} + (b_i)_{i=1}^{\infty} \in U$.

㊽ For any scalar $c$ and any element $(a_i)_{i=1}^{\infty} \in U$, we have $c(a_i)_{i=1}^{\infty} \in U$.

---

**Condition 1.**

The zero vector in $V$ is the zero sequence $(0) = (0, 0, 0, \ldots)$. Clearly, this sequence satisfies the recurrence relation $a_{k+2} - 5a_{k+1} + 3a_k = 0$. Thus the zero vector $(0) \in U$, hence condition 1 is met.

---

**Condition 2.**

To check condition 2, take $(a_i)_{i=1}^{\infty}, (b_i)_{i=1}^{\infty} \in U$.
We want to show that the sum

$$(a_i)_{i=1}^{\infty} + (b_i)_{i=1}^{\infty} = (a_i + b_i)_{i=1}^{\infty}$$

is also in $U$.

---

Since $(a_i)_{i=1}^{\infty}, (b_i)_{i=1}^{\infty} \in U$, these sequences satisfy the recurrence relation

$$a_{k+2} - 5a_{k+1} + 3a_k = 0 \tag{*}$$

and

$$b_{k+2} - 5b_{k+1} + 3b_k = 0$$

for $k = 1, 2, \ldots$.
Using these two relations, we have

$$\begin{aligned}&(a_{k+2} + b_{k+2}) - 5(a_{k+1} + b_{k+1}) + 3(a_k + b_k)\\ &= (a_{k+2} - 5a_{k+1} + 3a_k) + (b_{k+2} - 5b_{k+1} + 3b_k)\\ &= 0 + 0 = 0.\end{aligned}$$

Thus the sum $(a_i)_{i=1}^{\infty} + (b_i)_{i=1}^{\infty} = (a_i + b_i)_{i=1}^{\infty}$ satisfies the recurrence relation and it is also in $U$.
Hence condition 2 is met.

---

**Condition 3.**

To check condition 3, take $(a_i)_{i=1}^{\infty} \in U$ and let $c \in \mathbb{R}$ be a scalar.
Then the sequence $(a_i)_{i=1}^{\infty}$ satisfies (*). Multiplying (*) by $c$, we have

$$(ca_{k+2}) - 5(ca_{k+1}) + 3(ca_k) = 0.$$

This implies that the scalar product $c(a_i)_{i=1}^{\infty} = (ca_i)_{i=1}^{\infty}$ satisfies the recurrence relation, and hence it is in $U$. Thus condition 3 is satisfied.

We have checked all subspace criteria, and thus $U$ is a subspace of the vector space $V$.

(49) Let $V$ be a vector space over a field $K$. If $W_1$ and $W_2$ are subspaces of $V$, then prove that the subset

$$W_1 + W_2 := \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in W_1, \mathbf{y} \in W_2\}$$

is a subspace of the vector space $V$.

**Proof.**

We prove the following subspace criteria:

(50) The zero vector $\mathbf{0}$ of $V$ is in $W_1 + W_2$.

(51) For any $\mathbf{u}, \mathbf{v} \in W_1 + W_2$, we have $\mathbf{u} + \mathbf{v} \in W_1 + W_2$.

(52) For any $\mathbf{v} \in W_1 + W_2$ and $r \in K$, we have $r\mathbf{v} \in W_1 + W_2$.

Since $W_1$ and $W_2$ are subspaces of $V$, the zero vector $\mathbf{0}$ of $V$ is in both $W_1$ and $W_2$.
Thus we have

$$\mathbf{0} = \mathbf{0} + \mathbf{0} \in W_1 + W_2.$$

So condition 1 is met.

Next, let $\mathbf{u}, \mathbf{v} \in W_1 + W_2$.
Since $\mathbf{u} \in W_1 + W_2$, we can write

$$\mathbf{u} = \mathbf{x} + \mathbf{y}$$

for some $\mathbf{x} \in W_1$ and $\mathbf{y} \in W_2$.
Similarly, we write

$$\mathbf{v} = \mathbf{x}' + \mathbf{y}'$$

for some $\mathbf{x}' \in W_1$ and $\mathbf{y}' \in W_2$.

Then we have

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (\mathbf{x} + \mathbf{y}) + (\mathbf{x}' + \mathbf{y}') \\ &= (\mathbf{x} + \mathbf{x}') + (\mathbf{y} + \mathbf{y}'). \end{aligned}$$

Since $\mathbf{x}$ and $\mathbf{x}'$ are both in the vector space $W_1$, their sum $\mathbf{x} + \mathbf{x}'$ is also in $W_1$.
Similarly we have $\mathbf{y} + \mathbf{y}' \in W_2$ since $\mathbf{y}, \mathbf{y}' \in W_2$.

Thus from the expression above, we see that

$$\mathbf{u} + \mathbf{v} \in W_1 + W_2,$$

hence condition 2 is met.

Finally, let $\mathbf{v} \in W_1 + W_2$ and $r \in K$.
Then there exist $\mathbf{x} \in W_1$ and $\mathbf{y} \in W_2$ such that

$$\mathbf{v} = \mathbf{x} + \mathbf{y}.$$

Since $W_1$ is a subspace, it is closed under scalar multiplication. Hence we have $r\mathbf{x} \in W_1$.
Similarly, we have $r\mathbf{y} \in W_2$.

It follows from this observation that

$$\begin{aligned} r\mathbf{v} &= r(\mathbf{x} + \mathbf{y}) \\ &= r\mathbf{x} + r\mathbf{y} \in W_1 + W_2, \end{aligned}$$

and thus condition 3 is met.

---

Therefore, by the subspace criteria $W_1 + W_2$ is a subspace of $V$.

## Related Question.

Let $U$ and $V$ be finite dimensional subspaces in a vector space over a scalar field $K$.
Then prove that

$$\dim(U + V) \leq \dim(U) + \dim(V).$$

For a proof, see the post “Dimension of the sum of two subspaces“.

# Linearly Independency of General Vectors

**Definition**

Let $V$ be a vector space over a scalar field $K$. Let $= \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be a set of vectors in $V$.

① We say that $S$ is **linearly independent** if whenever $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ we have $c_1 = \cdots = c_k = 0$.

## Problems

① Let $C[3, 10]$ be the vector space consisting of all continuous functions defined on the interval $[3, 10]$. Consider the set $S = \{\sqrt{x}, x^2\}$ in $C[3, 10]$. Show that the set $S$ is linearly independent in $C[3, 10]$.

### Proof.

Note that the zero vector in $C[3, 10]$ is the zero function $\theta(x) = 0$.

Let us consider a linear combination

$$a_1\sqrt{x} + a_2x^2 = \theta(x) = 0.$$

We want to show that $a_1 = a_2 = 0$.

Since this equality holds for any value of $x$ between $3$ and $10$, letting $x = 4$ and $x = 9$ yields the system of linear equations

$$\begin{aligned} 2a_1 + 4a_2 &= 0 \\ 3a_1 + 81a_2 &= 0. \end{aligned}$$

Solving this system, we obtain $a_1 = a_2 = 0$, and hence the set $S$ is linearly independent.

### Comment.

We could have taken any two values between $3$ and $10$ for $x$ instead of $4$ and $9$, but the choice $x = 4$ and $x = 9$ made solving the system easier.

Also, we took two values for $x$ because we had two unknowns $a_1$, $a_2$, and thus we needed two equations to determine these unknown.

For a similar problem, show that the set $\{\sqrt{x}, x^2, x\}$ in the vector space $C[1, 10]$ is linearly independent.

In this case, there will be three unknowns that you want to show to be zero.

So you need to take three values for $x$.

A natural choice will be $x = 1, 4, 9$.

② Let $C[-\pi, \pi]$ be the vector space of all continuous functions defined on the interval $[-\pi, \pi]$. Show that the subset $\{\cos(x), \sin(x)\}$ in $C[-\pi, \pi]$ is linearly independent.

### Proof.

Note that the zero vector in the vector space $C[-\pi, \pi]$ is the zero function

$$\theta(x) := 0.$$

Let us consider a linear combination

$$a_1 \cos(x) + a_2 \sin(x) = \theta(x) = 0. \qquad (*)$$

If this linear combination has only the zero solution $a_1 = a_2 = 0$, then the set $\{\cos(x), \sin(x)\}$ is linearly independent.

---

The equality (*) should be true for any values of $x \in [-\pi, \pi]$.
Setting $x = 0$, we obtain from (*) that

$$a_1 = 0$$

since $\cos(0) = 1, \sin(0) = 0$.

We also set $x = \pi/2$ and we obtain

$$a_2 = 0$$

since $\cos(\pi/2) = 0, \sin(\pi/2) = 1$.

Therefore, we have $a_1 = a_2 = 0$ and we conclude that the set $\{\cos(x), \sin(x)\}$ is linearly independent.

3 Let $C[-2\pi, 2\pi]$ be the vector space of all continuous functions defined on the interval $[-2\pi, 2\pi]$. Consider the functions $f(x) = \sin^2(x)$ and $g(x) = \cos^2(x)$ in $C[-2\pi, 2\pi]$. Prove or disprove that the functions $f(x)$ and $g(x)$ are linearly independent.
*(The Ohio State University)*

### Proof.

To determine whether $f(x)$ and $g(x)$ are linearly independent or not, consider the linear combination

$$c_1 f(x) + c_2 g(x) = 0,$$

equivalently

$$c_1 \sin^2(x) + c_2 \cos^2(x) = 0, \qquad (*)$$

where $c_1, c_2$ are scalars.

If the only scalars satisfying the above equality are $c_1 = 0, c_2 = 0$, then $f(x)$ and $g(x)$ are linearly independent, otherwise they are linearly dependent.

Note that this is an equality as functions.
That is, this equality must hold for any $x$ in the interval $[-2\pi, 2\pi]$ .

Let $x = 0$ . Then as $\sin(0) = 0$ and $\cos(0) = 1$, we obtain $c_2 = 0$ from (*).
Next, let $x = \pi/2$ . Then as $\sin(\pi/2) = 1$ and $\cos(\pi/2) = 0$ , we obtain $c_1 = 0$ from (*).

Therefore, we must have $c_1 = c_2 = 0$, and hence the functions $f(x) = \sin^2(x)$ and $g(x) = \cos^2(x)$ are linearly independent.

## Comment.

This is one of the midterm 2 exam problems for Linear Algebra (Math 2568) in Autumn 2017.

Here is the most common mistake.
The linear combination $c_1 \sin^2(x) + c_2 \cos^2(x)$ is a function defined over the interval $[-2\pi, 2\pi]$ and we are assuming it is the zero function.

So saying that "if $c_1 = 1, c_2 = 0$, then $c_1 \sin^2(x) + c_2 \cos^2(x)$ is zero at $x = 0$ , hence $f(x)$ and $g(x)$ are linearly independent" is totally wrong.

What you are claiming here is that the function $\sin^2(x)$ is zero at $x = 0$ , hence it is the zero function.
This is clearly wrong as $\sin^2(x)$ is not the zero function.

**4** Let $f(x) = \sin^2(x)$, $g(x) = \cos^2(x)$, and $h(x) = 1$. These are vectors in $C[-1, 1]$. Determine whether the set $\{f(x),\ g(x),\ h(x)\}$ is linearly dependent or linearly independent.
(*The Ohio State University*)

## Solution.

We claim that the set is linearly dependent. To show the claim, we need to find nontrivial scalars $c_1, c_2, c_3$ such that

$$c_1 f(x) + c_2 g(x) + c_3 h(x) = 0.$$

From trigonometry, we know the identity

$$\sin^2(x) + \cos^2(x) = 1.$$

This implies that we have

$$\sin^2(x) + \cos^2(x) - 1 = 0.$$

So we can choose $c_1 = 1, c_2 = 1, c_3 = -1$ , and thus the set is linearly dependent.

**5** Let $c_1, c_2, \dots, c_n$ be mutually distinct real numbers. Show that exponential functions

$$e^{c_1 x}, e^{c_2 x}, \dots, e^{c_n x}$$

are linearly independent over $\mathbb{R}$.

## Hint.

⑥ Consider a linear combination

$$a_1 e^{c_1 x} + a_2 e^{c_2 x} + \cdots + a_n e^{c_n x} = 0.$$

⑦ Differentiate this equality $n - 1$ times and you will get $n$ equations.

⑧ Write a matrix equation for the system. You will see the Vandermonde matrix.

## Proof.

Suppose that we have a linear combination of these functions that is zero.

Namely, suppose we have

$$a_1 e^{c_1 x} + a_2 e^{c_2 x} + \cdots + a_n e^{c_n x} = 0$$

for some real numbers $a_1, a_2, \dots, a_n$ .

We want to show that the coefficients $a_1, a_2, \dots, a_n$ are all zero.

---

By differentiating the equation, we obtain

$$a_1 c_1 e^{c_1 x} + a_2 c_2 e^{c_2 x} + \cdots + a_n c_n e^{c_n x} = 0.$$

Differentiating repeatedly we further obtain the equalities

$$\begin{aligned}
&a_1 c_1^2 e^{c_1 x} + a_2 c_2^2 e^{c_2 x} + \cdots + a_n c_n^2 e^{c_n x} = 0 \\
&a_1 c_1^3 e^{c_1 x} + a_2 c_2^3 e^{c_2 x} + \cdots + a_n c_n^3 e^{c_n x} = 0 \\
&\dots \\
&a_1 c_1^{n-1} e^{c_1 x} + a_2 c_2^{n-1} e^{c_2 x} + \cdots + a_n c_n^{n-1} e^{c_n x} = 0
\end{aligned}$$

We rewrite these $n$ equations into the following matrix equation.

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ c_1 & c_2 & \dots & c_n \\ c_1^2 & c_2^2 & \dots & c_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ c_1^{n-1} & c_2^{n-1} & \dots & c_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 e^{c_1 x} \\ a_2 e^{c_2 x} \\ \vdots \\ a_n e^{c^n x} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{*}$$

---

The determinant of the left matrix is

$$\det \begin{bmatrix} 1 & 1 & \dots & 1 \\ c_1 & c_2 & \dots & c_n \\ c_1^2 & c_2^2 & \dots & c_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ c_1^{n-1} & c_2^{n-1} & \dots & c_n^{n-1} \end{bmatrix} = \prod_{i<j}(c_j - c_i)$$

by the Vandermonde determinant.

Since by assumption $c_1, \dots, c_n$ are distinct, the determinant is not zero.
Therefore by multiplying equality (*) by the inverse on the left, we obtain

$$\begin{bmatrix} a_1 e^{c_1 x} \\ a_2 e^{c_2 x} \\ \vdots \\ a_n e^{c^n x} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since the functions $e^{c_i x}$ are always positive, we must have $a_1 = a_2 = \cdots = a_n = 0$ as required.
Therefore the functions $e^{c_1 x}, \dots, e^{c_n x}$ are linearly independent.

## Comment.

The determinant that we considered above is called the Wronskian for the set of functions $\{e^{c_1 x}, e^{c_2 x}, \dots, e^{c_n x}\}$.

## Related Question.

The following problems are more concrete versions of the current problem.

**Problem.**
Let $C[-1,1]$ be the vector space over $\mathbb{R}$ of all continuous functions defined on the interval $[-1,1]$. Let

$$V := \{f(x) \in C[-1,1] \mid f(x) = ae^x + be^{2x} + ce^{3x}, a, b, c \in \mathbb{R}\}$$

be a subset in $C[-1,1]$.

**(a)** Prove that $V$ is a subspace of $C[-1,1]$.
**(b)** Prove that the set $B = \{e^x, e^{2x}, e^{3x}\}$ is a basis of $V$.
**(c)** Prove that

$$B' = \{e^x - 2e^{3x}, e^x + e^{2x} + 2e^{3x}, 3e^{2x} + e^{3x}\}$$

is a basis for $V$.

See the post ↴
Exponential Functions Form a Basis of a Vector Space
for the solution.

By calculating the Wronskian, determine whether the set of exponential functions

$$\{e^x, e^{2x}, e^{3x}\}$$

is linearly independent on the interval $[-1, 1]$.

The solutions is given in the post ↴

Using the Wronskian for Exponential Functions, Determine Whether the Set is Linearly Independent

⑨ Let $V$ be an $n$-dimensional vector space over a field $K$.
Suppose that $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent vectors in $V$. Are the following vectors linearly independent?

$$\mathbf{v}_1 + \mathbf{v}_2, \quad \mathbf{v}_2 + \mathbf{v}_3, \quad \dots, \quad \mathbf{v}_{k-1} + \mathbf{v}_k, \quad \mathbf{v}_k + \mathbf{v}_1.$$

If it is linearly dependent, give a non-trivial linear combination of these vectors summing up to the zero vector.

## Hint.

⑩ Consider a linar combination of these vectors summing up to the zero vector .

⑪ Check whether the coefficients are zero or not.

⑫ To do this, form a matrix equation.

## Solution.

Consider a linear combination

$$c_1(\mathrm{v}_1 + \mathrm{v}_2) + c_2(\mathrm{v}_2 + \mathrm{v}_3) + \cdots + c_{k-1}(\mathrm{v}_{k-1} + \mathrm{v}_k) + c_k(\mathrm{v}_k + \mathrm{v}_1) = 0.$$

Collecting similar terms, we obtain

$$(c_1 + c_k)\mathrm{v}_1 + (c_1 + c_2)\mathrm{v}_2 + \cdots + (c_{k-2} + c_{k-1})\mathrm{v}_{k-1} + (c_{k-1} + c_k)\mathrm{v}_k = 0.$$

Since $\mathrm{v}_1, \mathrm{v}_2, \dots, \mathrm{v}_k$ are linearly independent, we have

$$\begin{aligned} c_1 + c_k &= 0 \\ c_1 + c_2 &= 0 \\ &\vdots \\ c_{k-2} + c_{k-1} &= 0 \\ c_{k-1} + c_k &= 0 \end{aligned}$$

This is equivalent to the following matrix equality

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 1 \\ 1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{k-1} \\ c_k \end{bmatrix} = 0.$$

We reduce the coefficient matrix using elementary row operations and obtain the matrix

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 & -1 \\ 0 & 0 & 1 & \dots & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & (-1)^k \\ 0 & 0 & 0 & \dots & 0 & 0 & 1-(-1)^k. \end{bmatrix}$$

Thus if $k$ is odd, then the system has only the trivial solution $c_i = 0$.

On the other hand, if $k$ is even, then the matrix is singular. Hence we have a non-trivial solution.

For example, letting $c_k = 1$, we see that

$$c_1 = -1, c_2 = 1, c_3 = -1, c_4 = 1, \dots, c_{k-1} = 1, c_k = 1$$

is a non-trivial solution.

Thus we have when $k$ is even

$$-(\mathrm{v}_1 + \mathrm{v}_2) + (\mathrm{v}_2 + \mathrm{v}_3) - \cdots + (\mathrm{v}_{k-1} + \mathrm{v}_k) - (\mathrm{v}_k + \mathrm{v}_1) = 0.$$

### Comment.

We used elementary row operations in the solution because we wanted to have a nontrivial linear combination.
From the computation, the determinant of the coefficient matrix can be computed easily.

The other method to calculate the determinant is to use the cofactor expansion. See problem Calculate determinants of matrices for details.

⑬ Let $V$ be a vector space over a scalar field $K$. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be the set of vectors in $V$, where $n \geq 2$. Then prove that the set $S$ is linearly dependent if and only if at least one of the vectors in $S$ can be written as a linear combination of remaining vectors in $S$.

### Proof.

$(\implies)$: If $S$ is linearly dependent, then a vector is a linear combination

Suppose that the set $S = \{\mathrm{v}_1, \mathrm{v}_2, \dots, \mathrm{v}_n\}$ is linearly dependent. Then there exists scalars $c_1, c_2, \dots, c_n$ in $K$ such that

$$c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + \cdots + c_n\mathrm{v}_n = 0$$

and not all of $c_1, \dots, c_n$ are zero.
(This is just the definition of linearly dependency.)

Since at least one of $c_1, \dots, c_n$ is nonzero, without loss of generality we may assume that $c_1 \neq 0$.
Then we have

$$c_1\mathrm{v}_1 = -c_2\mathrm{v}_2 - c_3\mathrm{v}_3 - \cdots - c_n\mathrm{v}_n$$

and we divide this by $c_1 \neq 0$ and obtain

$$\mathrm{v}_1 = -\frac{c_2}{c_1}\mathrm{v}_2 - \frac{c_3}{c_1}\mathrm{v}_3 - \cdots - \frac{c_n}{c_1}\mathrm{v}_n.$$

Therefore, we have written the vector $\mathrm{v}_1$ as a linear combination of $\mathrm{v}_2, \mathrm{v}_3, \dots, \mathrm{v}_n$.

### $(\Longleftarrow)$: If a vector is a linear combination, then $S$ is linearly dependent

Suppose now that some vector of $S$ is a linear combination of the remaining vectors.
Without loss of generality, we may assume that $\mathrm{v}_1$ is a linear combination of the remaning vectors. So we have

$$\mathrm{v}_1 = c_2\mathrm{v}_2 + c_3\mathrm{v}_3 + \cdots + c_n\mathrm{v}_n,$$

for some scalars $c_2, c_3, \dots, c_n$ in $K$.

This implies that we have

$$\mathrm{v}_1 - c_2\mathrm{v}_2 - c_3\mathrm{v}_3 - \cdots - c_n\mathrm{v}_n = 0,$$

and this is a nontrivial linear combination of vectors in $S$ because the coefficient of $\mathrm{v}_1$ is $1 \neq 0$.
Thus, the set $S$ is linearly dependent.

⑭ Let $V$ be a vector space over a field $K$. Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be linearly independent vectors in $V$. Let $U$ be the subspace of $V$ spanned by these vectors, that is, $U = \mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$. Let $\mathbf{u}_{n+1} \in V$. Show that $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{u}_{n+1}$ are linearly independent if and only if $\mathbf{u}_{n+1} \notin U$.

### Proof.

$(\Longrightarrow)$

Suppose that the vectors $\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_n, \mathrm{u}_{n+1}$ are linearly independent. If $\mathrm{u}_{n+1} \in U$, then $\mathrm{u}_{n+1}$ is a linear combination of $\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_n$.
Thus, we have

$$\mathrm{u}_{n+1} = c_1\mathrm{u}_1 + c_2\mathrm{u}_2 + \cdots + c_n\mathrm{u}_n$$

for some scalars $c_1, c_2, \ldots, c_n \in K$ .

However, this implies that we have a nontrivial linear combination

$$c_1\mathrm{u}_1 + c_2\mathrm{u}_2 + \cdots + c_n\mathrm{u}_n - \mathrm{u}_{n+1} = 0.$$

This contradicts that $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n, \mathrm{u}_{n+1}$ are linearly independent. Hence $\mathrm{u}_{n+1} \notin U$.

---

( $\Longleftarrow$ ) Suppose now that $\mathrm{u}_{n+1} \notin U$.
If the vectors $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n, \mathrm{u}_{n+1}$ are linearly dependent, then there exists $c_1, c_2 \ldots, c_n, c_{n+1} \in K$ such that
not all of them are zero and

$$c_1\mathrm{u}_1 + c_2\mathrm{u}_2 + \cdots + c_n\mathrm{u}_n + c_{n+1}\mathrm{u}_{n+1} = 0.$$

We claim that $c_{n+1} \neq 0$. If $c_{n+1} = 0$, then we have

$$c_1\mathrm{u}_1 + c_2\mathrm{u}_2 + \cdots + c_n\mathrm{u}_n = 0$$

and since $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n$ are linearly independent, we must have $c_1 = c_2 = \cdots = c_n = 0$. This means that all $c_i$ are zero but this contradicts our choice of $c_i$ . Thus $c_{n+1} \neq 0$.

Then we have

$$\mathrm{u}_{n+1} = \frac{-c_1}{c_{n+1}}\mathrm{u}_1 + \frac{-c_2}{c_{n+1}}\mathrm{u}_2 + \cdots + \frac{-c_n}{c_{n+1}}\mathrm{u}_n .$$

(Note: we needed to check $c_{n+1} \neq 0$ to divide by it.)

This implies that $\mathrm{u}_{n+1}$ is a linear combination of vectors $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n$ , and thus $\mathrm{u}_{n+1} \in U$, a contradiction.
Therefore, the vectors $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n, \mathrm{u}_{n+1}$ are linearly independent.

⑮ By calculating the **Wronskian**, determine whether the set of exponential functions $\{e^x, e^{2x}, e^{3x}\}$ is linearly independent on the interval $[-1, 1]$.

### Solution.

The Wronskian for the set $\{e^x, e^{2x}, e^{3x}\}$ is given by

$$W(x) := \begin{vmatrix} e^x & e^{2x} & e^{3x} \\ e^x & 2e^{2x} & 3e^{3x} \\ e^x & 4e^{2x} & 9e^{3x} \end{vmatrix}.$$

We compute

$$W(x) = e^x e^{2x} e^{3x} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{vmatrix}$$

$$= 2e^{6x}.$$

Since the Wronskian $W(x) = 2e^{6x}$ is never zero, we conclude that the set $\{e^x, e^{2x}, e^{3x}\}$ is linearly independent.

---

(Note that in general, we just need to show that $W(x_0) \neq 0$ for some point $x_0 \in [0, 1]$.

For example, since we have $W(0) = 2 \neq 0$, we can conclude that $\{e^x, e^{2x}, e^{3x}\}$ is linearly independent.)

## Related Question.

Try the following generalized version.

Let $c_1, c_2, \dots, c_n$ be mutually distinct real numbers.

Show that functions

$$e^{c_1 x}, e^{c_2 x}, \dots, e^{c_n x}$$

are linearly independent over $\mathbb{R}$.

The solution is given by Exponential Functions are Linearly Independent.

# Bases and Coordinate Vectors

**Definition**

Let $V$ be a vector space over a scalar field $K$.

1. A basis $B$ for $V$ is a linearly independent spanning set for $V$.
2. Suppose that $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for $V$. Let $\mathbf{v} \in V$ and write it as $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, where $c_1, \dots, c_n \in K$. Then the coordinate vector of $\mathbf{v}$ with respect to the basis $B$ is

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n.$$

**Summary**

Let $V$ be a vector space over a scalar field $K$. Suppose that $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for $V$. Let $S = \{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ be a set of vectors in $V$. Let $T = \{[\mathbf{w}_1]_B, \dots, [\mathbf{w}_k]_B\}$ be the set of the coordinate vectors of $S$.

1. Any vector $\mathbf{v}$ can be uniquely written as $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, where $c_1, \dots, c_n \in K$.
2. The dimension of the coordinate vector $[\mathbf{v}]_B$ is the dimension of the vector space $V$.
3. The set $S$ is linearly independent if and only if $T$ is linearly independent.
4. $S$ is a basis for $\text{Span}(S)$ if and only if $T$ is a basis for $\text{Span}(T)$ in $\mathbb{R}^n$.

## Problems

1. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a basis for a vector space $V$ over a scalar field $K$. Then show that any vector $\mathbf{v} \in V$ can be written uniquely as

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3,$$

where $c_1, c_2, c_3$ are scalars.

### Proof.

Since $B$ is a basis for $V$, any vector $\mathbf{v} \in V$ is a linear combination of basis vectors in $B$. Thus, there exist scalars $c_1, c_2, c_3 \in K$ such that

$$\mathbf{v} = c_1\mathbf{v}_{+}c_2\mathbf{v}_2 + c_3\mathbf{v}_3.$$

Hence such an expression as a linear combination of basis vectors exists.

We now show that such representation of $\mathbf{v}$ is unique.
Suppose we have another representaion

$$\mathbf{v} = d_1\mathbf{v}_{+}d_2\mathbf{v}_2 + d_3\mathbf{v}_3$$

for some scalars $d_1, d_2, d_3 \in K$.

Then we have

$$c_1\mathbf{v}_{+}c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{v} = d_1\mathbf{v}_{+}d_2\mathbf{v}_2 + d_3\mathbf{v}_3.$$

Thus, we obtain

$$(c_1 - d_1)\mathbf{v}_{+}(c_2 - d_2)\mathbf{v}_2 + (c_3 - d_3)\mathbf{v}_3 = \mathbf{0}.$$

Since $B$ is a basis, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.

Hence the coefficients of the above linear combination must be all zero, and thus we obtain

$$c_1 - d_1 = 0, \quad c_2 - d_2 = 0, \quad c_3 - d_3 = 0,$$

equivalently, we have

$$c_1 = d_1, \quad c_2 = d_2, \quad c_3 = d_3.$$

Therefore two representations of the vector $\mathbf{v}$ are the same, and thus the representation of $\mathbf{v}$ as a linear combination of basis vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is unique.

② Show that the set

$$S = \{1, 1 - x, 3 + 4x + x^2\}$$

is a basis of the vector space $P_2$ of all polynomials of degree $2$ or less.

**Proof.**

We know that the set $B = \{1, x, x^2\}$ is a basis for the vector space $P_2$.
With respect to this basis $B$, the coordinate vectors of vectors in $S$ are

$$[1]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [1 - x]_B = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad [3 + 4x + x^2]_B = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}.$$

Recall the fact that the set $S$ is a basis for $P_2$ if and only if the coordinate vectors $T = \{[1]_B, [1-x]_B, [3+4x+x^2]_B\}$ is a basis for $R^3$.

Hence we check that $T$ is a basis for $R^3$. Note that $T$ is a linearly independent set. In fact, if we have

$$a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} = \mathbf{0},$$

then we have

$$\begin{bmatrix} a_1 + a_2 + 3a_3 \\ -a_2 + 4a_3 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and we obtain $a_1 = a_2 = a_3 = 0$, and thus $T$ is a linearly independent set.

Since $T$ consists of three linearly independent vectors in $\mathbb{R}^3$ and the dimension of $\mathbb{R}^3$ is $3$, the set $T$ must be a basis for $\mathbb{R}^3$.

Therefore, the set $S$ is a basis for $P_2$.

③ Let $V$ be the vector space of all $2 \times 2$ matrices, and let the subset $S$ of $V$ be defined by $S = \{A_1, A_2, A_3, A_4\}$, where

$$A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 4 \end{bmatrix}, \quad A_3 = \begin{bmatrix} -1 & 0 \\ 1 & -10 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 3 & 7 \\ -2 & 6 \end{bmatrix}.$$

Find a basis of the span $\mathrm{Span}(S)$ consisting of vectors in $S$ and find the dimension of $\mathrm{Span}(S)$.

## Proof.

Let $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ be the standard basis for the vector space $V$, where

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

With respect to the basis $B$, we find the coordinate vectors for $A_1, A_2, A_3, A_4$ as follows.

Since we have

$$A_1 = E_{11} + 2E_{12} - E_{21} + 3E_{22},$$

the coordinate vector for $A_1$ is

$$[A_1]_B = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}.$$

Similarly, we have

$$[A_2]_B = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 4 \end{bmatrix}, [A_3]_B = \begin{bmatrix} -1 \\ 0 \\ 1 \\ -10 \end{bmatrix}, [A_4]_B = \begin{bmatrix} 3 \\ 7 \\ -2 \\ 6 \end{bmatrix}.$$

To find a basis for $\mathrm{Span}(S)$ among vectors in $S$, we first find a basis for $\mathrm{Span}(T)$ among vectors in

$$T = \{[A_1]_B, [A_2]_B, [A_3]_B, [A_4]_B\}.$$

Let form a matrix whose columns are vectors in $T$. That is,

$$\begin{bmatrix} 1 & 0 & -1 & 3 \\ 2 & -1 & 0 & 7 \\ -1 & 1 & 1 & -2 \\ 3 & 4 & -10 & 6 \end{bmatrix}.$$

We apply the elementary row operations as follows and obtain a reduced row echelon form matrix.

$$\begin{bmatrix} 1 & 0 & -1 & 3 \\ 2 & -1 & 0 & 7 \\ -1 & 1 & 1 & -2 \\ 3 & 4 & -7 & 6 \end{bmatrix} \xrightarrow{\substack{R_2-2R_1 \\ R_3+R_1 \\ R_4-3R_1}} \begin{bmatrix} 1 & 0 & -1 & 3 \\ 0 & -1 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 4 & -7 & -3 \end{bmatrix} \xrightarrow{\substack{R_3+R_2 \\ R_4+4R_2}} \begin{bmatrix} 1 & 0 & -1 & 3 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow{\substack{R_1+R_4 \\ R_2-2R_4 \\ R_3-2R_4}} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{1R_2 \\ R_3 \leftrightarrow R_4}} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The the first three columns of the reduced row echelon form contains the leading 1's.
Thus by, what I call, the leading $1$ method, it follows that

$$\{[A_1]_B, [A_2]_B, [A_3]_B\}$$

is a basis for $\mathrm{Span}(T)$.

Therefore, by the correspondence of coordinate vectors, we obtain that

$$\{A_1, A_2, A_3\}$$

is a basis of $\mathrm{Span}(S)$.

Since the basis $\{A_1, A_2, A_3\}$ consists of three vectors, the dimension of $\mathrm{Span}(S)$ is $3$.

④ Let $P_2$ be the vector space of all polynomials of degree $2$ or less with real coefficients. Let

$$S = \{1 + x + 2x^2, \quad x + 2x^2, \quad -1, \quad x^2\}$$

be the set of four vectors in $P_2$. Then find a basis of the subspace $\mathrm{Span}(S)$ among the vectors in $S$.
(*The Ohio State University*)

---

## Solution.

Let $B = \{1, x, x^2\}$ be the standard basis of the vector space $P_2$.
With respect to the basis $B$, the coordinate vectors of vectors in $S$ are

$$[1 + x + 2x^2]_B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad [x + 2x^2]_B = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad [-1]_B = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \quad [x^2]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Let

$$T = \left\{ [1 + x + 2x^2]_B, [x + 2x^2]_B, [-1]_B, [x^2]_B \right\}$$

be the set of these coordinate vectors.

We then find a basis of $\mathrm{Span}(T)$ among vectors in $T$ by the leading 1 method.
We reduce the augmented matrix by elementary row operations as follows. We have

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The last matrix is in reduced row echelon form and the first, the second, and the fourth columns contain the leading 1's.
Therefore, it follows that

$$\left\{ [1 + x + 2x^2]_B, [x + 2x^2]_B, [x^2]_B \right\}$$

is a basis of $\mathrm{Span}(T)$, and hence

$$\{1 + x + 2x^2, \quad x + 2x^2, \quad x^2\}$$

is a basis of $\mathrm{Span}(S)$ consisting of the vectors of $S$.

⑤ Let $P_2$ be the vector space of all polynomials of degree two or less.
Consider the subset in $P_2$

$$Q = \{p_1(x), p_2(x), p_3(x), p_4(x)\},$$

where

$$\begin{aligned} p_1(x) &= x^2 + 2x + 1, & p_2(x) &= 2x^2 + 3x + 1, \\ p_3(x) &= 2x^2, & p_4(x) &= 2x^2 + x + 1. \end{aligned}$$

**(a)** Use the basis $B = \{1, x, x^2\}$ of $P_2$, give the coordinate vectors of the vectors in $Q$.

**(b)** Find a basis of the span $\mathrm{Span}(Q)$ consisting of vectors in $Q$.
**(c)** For each vector in $Q$ which is not a basis vector you obtained in (b), express the vector as a linear combination of basis vectors.

## Solution.

### Coordinate vectors

**(a)** With respect to the given basis $B = \{1, x, x^2\}$, the coordinate vectors are

$$[p_1(x)]_B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \qquad [p_2(x)]_B = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix},$$
$$[p_3(x)]_B = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \qquad [p_4(x)]_B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

For example, to obtain $[p_3(x)]_B$, we express $p_3(x)$ as a linear combination of vectors in $B$ and we get

$$p_3(x) = 0 \cdot 1 + 0 \cdot x + 2 \cdot x^2.$$

Note that the order of the sum in the linear combination is the order of basis vectors in $B = \{1, x, x^2\}$. Then we read the coefficients of this linear combination and obtain the coordinate vectors

$$[p_3(x)]_B = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

The rest are the same.

### Find a basis of the span

**(b)** Let

$$T := \{[p_1(x)]_B, [p_2(x)]_B, [p_3(x)]_B, [p_4(x)]_B\}$$
$$= \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

Then a basis of $\mathrm{Span}(Q)$ consisting of vectors in $Q$ corresponds to a basis of $\mathrm{Span}(T)$ consisting of vectors in $T$.

Consider the matrix whose columns are the vectors in $T$. We reduce it by elementary row operations as follows.

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 2 & 3 & 0 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} \xrightarrow[R_3-R_1]{R_2-2R_1} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 1 \end{bmatrix} \xrightarrow[R_3-R_2]{R_1-R_2}$$
$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 2 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2}R_3} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The last matrix is in reduced row echelon form and the first three columns contain the leading 1's.
Therefore the set consisting of the first three vectors in $T$

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \right\}$$

is a basis for $\mathrm{Span}(T)$.

Therefore by the coordinate vector correspondence, it follows that the set

$$\{p_1(x), p_2(x), p_3(x)\}$$

is a basis for $\mathrm{Span}(Q)$ consisting of vectors in $Q$.

### Linear combination of basis vectors

**(c)** From part (b), the only vector in $Q$ which is not a basis vector is $p_4(x)$.
Thus we want to express $p_4(x)$ as a linear combination of $p_1(x), p_2(x), p_3(x)$.
Consider a linear combination

$$a_1p_1(x) + a_2p_2(x) + a_3p_3(x) + a_4p_4(x) = 0. \tag{*}$$

By taking coordinate vectors, we obtain

$$a_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} + a_4 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 0.$$

The augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 2 & 3 & 0 & 1 & 0 \\ 1 & 2 & 2 & 2 & 0 \end{array}\right].$$

Note that this is the matrix we considered in part (b) except for the last zero column.
Thus by the same elementary row operations, we obtain the reduced row echelon form of this matrix

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right].$$

Therefore we obtain the solutions

$$\begin{aligned} a_1 &= -2a_4 \\ a_2 &= a_4 \\ a_3 &= -a_4 \end{aligned}$$

and $a_4$ is a free variable.

(This is the Gauss-Jordan elimination method.)

---

Let us take $a_4 = 1$. Then from the above solutions, we obtain $a_1 = -2, a_2 = 1, a_3 = -1$. We plug these values in the linear combination (*) and obtain

$$-2p_1(x) + p_2(x) - p_3(x) + p_4(x) = 0.$$

Solving this for $p_4(x)$, we obtain the linear combination

$$p_4(x) = 2p_1(x) - p_2(x) + p_3(x).$$

(Remark: the coefficients in the above linear combination of $p_4(x)$ is the fourth column of the reduced row echelon form.)

6 Let $P_2$ be the vector space of all polynomials with real coefficients of degree $2$ or less. Let $S = \{p_1(x), p_2(x), p_3(x), p_4(x)\}$, where

$$\begin{aligned} p_1(x) &= -1 + x + 2x^2, \quad & p_2(x) &= x + 3x^2 \\ p_3(x) &= 1 + 2x + 8x^2, \quad & p_4(x) &= 1 + x + x^2. \end{aligned}$$

**(a)** Find a basis of $P_2$ among the vectors of $S$.

**(b)** Let $B'$ be the basis you obtained in part (a). For each vector of $S$ which is not in $B'$, find the coordinate vector of it with respect to the basis $B'$.

(*The Ohio State University*)

## Solution.

### (a) Find a basis of $P_2$ among the vectors of $S$.

The vector space $P_2$ has a basis $B = \{1, x, x^2\}$ and the dimension of $P_2$ is $3$.

The coordinate vectors with respect to this basis are

$$[p_1(x)]_B = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad [p_2(x)]_B = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

$$[p_3(x)]_B = \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix}, \quad [p_4(x)]_B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Consider the vector space $\mathrm{Span}(T)$ spanned by

$$T = \{[p_1(x)]_B, [p_2(x)]_B, [p_3(x)]_B, [p_4(x)]_B\}.$$

We determine a basis vectors of $\mathrm{Span}(T)$ among vectors in $T$.
Consider the matrix $A$ whose column vectors are these coordinate vectors:

$$A = \begin{bmatrix} -1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 3 & 8 & 1 \end{bmatrix}.$$

Applying the elementary row operations, we obtain

$$A = \begin{bmatrix} -1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 3 & 8 & 1 \end{bmatrix} \xrightarrow{\substack{R_2+R_1 \\ R_3+2R_1}} \begin{bmatrix} -1 & 0 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 3 & 10 & 3 \end{bmatrix}$$

$$\xrightarrow[R_3-3R_2]{-R_1} \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & -3 \end{bmatrix} \xrightarrow{\substack{R_1+R_3 \\ R_2-3R_3}} \begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & -3 \end{bmatrix}.$$

The first three columns contain the leading 1's.
By the leading 1 method, we know that

$$\{[p_1(x)]_B, [p_2(x)]_B, [p_3(x)]_B\}$$

is a basis of the vector space $\mathrm{Span}(T)$.

It follows by the correspondence theorem that

$$\{p_1(x), p_2(x), p_3(x)\}$$

is a basis of $\mathrm{Span}(S)$.

This yields that $\mathrm{Span}(S)$ is a three dimensional subspace of $P_2$, which is also three dimensional.
Hence we have $P_2 = \mathrm{Span}(S)$.
Therefore

$$\{p_1(x), p_2(x), p_3(x)\}$$

is a basis of $P_2$.

> (b) For each vector of $S$ which is not in $B'$, find the coordinate vector of it with respect to the basis $B'$.

In part (a), we found a basis

$$B' = \{p_1(x), p_2(x), p_3(x)\}$$

of the vector space $P_2$.
So, we need to find the coordinate vector of $p_4(x)$ with respect to $B'$.

Thus, we need to solve

$$ap_1(x) + bp_2(x) + cp_3(x) = p_4(x)$$

for $a, b, c$.

Using the coordinate vectors, this is equivalent to solve

$$\begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 3 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

The augmented matrix of this system is exactly the matrix $A$ we dealt in part (a).
By the reduction of $A$ obtained in part (a) yields that

$$a = -4, b = 11, c = -3.$$

Hence we obtain

$$p_4(x) = -4p_1(x) + 11p_2(x) + 3p_3(x),$$

and thus the coordinate vector of $p_4(x)$ with respect to $B'$ is

$$[p_4(x)]_{B'} = \begin{bmatrix} -4 \\ 11 \\ 3 \end{bmatrix}.$$

Remark that this is the last column vector of the reduced row echelon form matrix of $A$.

7 Let $P_2$ be the vector space of all polynomials of degree two or less. Consider the subset in $P_2$

$$Q = \{p_1(x), p_2(x), p_3(x), p_4(x)\},$$

where

$$\begin{aligned} p_1(x) &= 1, & p_2(x) &= x^2 + x + 1, \\ p_3(x) &= 2x^2, & p_4(x) &= x^2 - x + 1. \end{aligned}$$

**(a)** Use the basis $B = \{1, x, x^2\}$ of $P_2$, give the coordinate vectors of the vectors in $Q$.
**(b)** Find a basis of the span $\mathrm{Span}(Q)$ consisting of vectors in $Q$.
**(c)** For each vector in $Q$ which is not a basis vector you obtained in (b), express the vector as a linear combination of basis vectors.
(*The Ohio State University*)

## Solution.

### (a) Use the basis $B=\{1,x,x^2\}$ of $P_2$, give the coordinate vectors of the vectors in $Q$

To obtain the coordinate vectors with respect to the given basis $B=\{1,x,x^2\}$, we fist express a given vector as a linear combination of basis vectors. ( You need to take care of the order of the basis vectors.)

For example, for the vector $p_4(x)$, we have the linear combination

$$p_4(x)=1\cdot 1+(-1)\cdot x+1\cdot x^2.$$

Thus reading the coefficient of the basis vectors $1,x,x^2$, we see that the coordinate vector of $p_4(x)$ with respect to the basis $B$ is

$$[p_4(x)]_B=\begin{bmatrix}1\\-1\\1\end{bmatrix}.$$

By the similar argument, the coordinate vectors are

$$\begin{aligned}
[p_1(x)]_B&=\begin{bmatrix}1\\0\\0\end{bmatrix}, & [p_2(x)]_B&=\begin{bmatrix}1\\1\\1\end{bmatrix},\\
[p_3(x)]_B&=\begin{bmatrix}0\\0\\2\end{bmatrix}, & [p_4(x)]_B&=\begin{bmatrix}1\\-1\\1\end{bmatrix}.
\end{aligned}$$

### (b) Find a basis of the span $\operatorname{Span}(Q)$ consisting of vectors in $Q$

Let

$$\begin{aligned}
T:&=\{[p_1(x)]_B,[p_2(x)]_B,[p_3(x)]_B,[p_4(x)]_B\}\\
&=\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\0\\2\end{bmatrix},\begin{bmatrix}1\\-1\\1\end{bmatrix}\right\}.
\end{aligned}$$

Then a basis of $\operatorname{Span}(Q)$ consisting of vectors in $Q$ corresponds to a basis of $\operatorname{Span}(T)$ consisting of vectors in $T$.

Consider the matrix whose columns are the vectors in $T$. We reduce it by elementary row operations as follows.

$$\begin{bmatrix}1&1&0&1\\0&1&0&-1\\0&1&2&1\end{bmatrix}\xrightarrow[R_3-R_2]{R_1-R_2}\begin{bmatrix}1&0&0&2\\0&1&0&-1\\0&0&2&2\end{bmatrix}\xrightarrow{\frac{1}{2}R_3}\begin{bmatrix}1&0&0&2\\0&1&0&-1\\0&0&1&1\end{bmatrix}.$$

The last matrix is in reduced row echelon form and the first three columns contain the leading 1's. Therefore the set consisting of the first three vectors in $T$

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \right\}$$

is a basis for $\mathrm{Span}(T)$.
(This is what I call the leading 1 method.)

Therefore by the coordinate vector correspondence, it follows that the set

$$\{p_1(x), p_2(x), p_3(x)\}$$

is a basis for $\mathrm{Span}(Q)$ consisting of vectors in $Q$.

> (c) For each vector in $Q$ which is not a basis vector you obtained in (b), express the vector as a linear combination of basis vectors

From the solution of part (b), we see that the vector $p_4(x)$ is the only vector in $Q$ which is not a basis vector.
Thus we want to express $p_4(x)$ as a linear combination of $p_1(x), p_2(x), p_3(x)$.

Consider a linear combination

$$a_1 p_1(x) + a_2 p_2(x) + a_3 p_3(x) + a_4 p_4(x) = 0. \tag{*}$$

By taking coordinate vectors, we obtain

$$a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} + a_4 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The augmented matrix is

$$\left[\begin{array}{cccr|c} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 1 & 0 \end{array}\right].$$

Note that this is the matrix we considered in part (b) except for the last zero column.
Thus by the same elementary row operations, we obtain the reduced row echelon form of this matrix

$$\left[\begin{array}{cccr|c} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right].$$

Therefore we obtain the solutions

$$\begin{aligned} a_1 &= -2a_4 \\ a_2 &= a_4 \\ a_3 &= -a_4 \end{aligned}$$

and $a_4$ is a free variable.

Setting $a_4 = 1$ we obtain $a_1 = -2, a_2 = 1, a_3 = -1$ from the above solutions.
We plug these values in the linear combination (*) and obtain

$$-2p_1(x)+p_2(x)-p_3(x)+p_4(x)=0.$$

Solving this for $p_4(x)$, the vector $p_4(x)$ can be expressed as

$$p_4(x)=2p_1(x)-p_2(x)+p_3(x).$$

8 Let $P_2$ be the vector space over $\mathbb{R}$ of all polynomials of degree $2$ or less.
Let $S=\{p_1(x),p_2(x),p_3(x)\}$, where

$$p_1(x)=x^2+1, \quad p_2(x)=6x^2+x+2, \quad p_3(x)=3x^2+x.$$

**(a)** Use the basis $B=\{x^2,x,1\}$ of $P_2$ to prove that the set $S$ is a basis for $P_2$.
**(b)** Find the coordinate vector of $p(x)=x^2+2x+3\in P_2$ with respect to the basis $S$.

## Solution.

### (a) Prove that the set $S$ is a basis for $P_2$.

The coordinate vectors of $p_1(x)$ with respect to the basis $B=\{x^2,x,1\}$ is given by

$$[p_1(x)]_B=\begin{bmatrix}1\\0\\1\end{bmatrix}$$

as $p_(x)$ can be written as a linear combination $p_1(x)=1\cdot x^2+0\cdot x+1\cdot 1$ of the basis vectors in $B$
.

Similarly, we have

$$[p_2(x)]_B=\begin{bmatrix}6\\1\\2\end{bmatrix} \text{ and } [p_3(x)]_B=\begin{bmatrix}3\\1\\0\end{bmatrix}.$$

Let $T=\{[p_1(x)]_B,[p_2(x)]_B,[p_3(x)]_B\}$.
Then $S$ is a basis for $P_2$ if and only if $T$ is a basis for $\mathbb{R}^3$.
Thus it remains to show that $T$ is a basis for $\mathbb{R}^3$.

---

Consider the matrix whose column vectors are vectors in $T$.

We have

$$\begin{bmatrix}1&6&3\\0&1&1\\1&2&0\end{bmatrix}\xrightarrow{R_3-R_1}\begin{bmatrix}1&6&3\\0&1&1\\0&-4&-3\end{bmatrix}\xrightarrow{\substack{R_1-6R_2\\R_3+4R_2}}$$
$$\begin{bmatrix}1&0&-3\\0&1&1\\0&0&1\end{bmatrix}\xrightarrow{\substack{R_1+3R_3\\R_2-R_3}}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

It follows that $T$ is linearly independent.

As $T$ consists of three linearly independent vectors in $\mathbb{R}^3$, it is a basis for $\mathbb{R}^3$.

Hence $S$ is a basis for $P_2$.

> (b) Find the coordinate vector of $p(x) = x^2 + 2x + 3 \in P_2$ with respect to the basis $S$.

To find the coordinate vector $[p(x)]_S$ with respect to the basis $S$, we express $p(x)$ as a linear combination of the basis vectors in $S$.

Thus we want to find scalars $c_1, c_2, c_3$ such that

$$p(x) = c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x).$$

Considering the coordinate vectors of both sides, this is equivalent to

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}.$$

We write this equation as a matrix equation

$$\begin{bmatrix} 1 & 6 & 3 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

---

To solve this, we apply elementary row operations to the augmented matrix as follows.

$$\left[\begin{array}{ccc|c} 1 & 6 & 3 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 2 & 0 & 3 \end{array}\right] \xrightarrow{R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 6 & 3 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & -4 & -3 & 2 \end{array}\right] \xrightarrow[R_3+4R_2]{R_1-6R_2}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & -3 & -11 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 10 \end{array}\right] \xrightarrow[R_2-R_3]{R_1+3R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 19 \\ 0 & 1 & 0 & -8 \\ 0 & 0 & 1 & 10 \end{array}\right].$$

(Note that elementary row operations are exactly the same as before.)

Hence the solution is $c_1 = 19, c_2 = -8, c_3 = 10$.

Thus, we have the linear combination

$$p(x) = 19p_1(x) - 8p_2(x) + 10p_3(x)$$

and the coordinate vector of $p(x)$ with respect to the basis $S$ is

$$[p(x)]_S = \begin{bmatrix} 19 \\ -8 \\ 10 \end{bmatrix}.$$

9 Let $\mathcal{P}_3$ be the vector space of all polynomials of degree $3$ or less.

Let

$$S = \{p_1(x), p_2(x), p_3(x), p_4(x)\},$$

where

$$\begin{array}{ll} p_1(x) = 1 + 3x + 2x^2 - x^3 & p_2(x) = x + x^3 \\ p_3(x) = x + x^2 - x^3 & p_4(x) = 3 + 8x + 8x^3. \end{array}$$

**(a)** Find a basis $Q$ of the span $\operatorname{Span}(S)$ consisting of polynomials in $S$.
**(b)** For each polynomial in $S$ that is not in $Q$, find the coordinate vector with respect to the basis $Q$.
*(The Ohio State University)*

## Solution.

### (a) Find a basis $Q$ of the span $\operatorname{Span}(S)$ consisting of polynomials in $S$.

Let $B = \{1, x, x^2, x^3\}$ be the standard basis for $\mathcal{P}_3$.

With respect to the basis $B$, the coordinate vectors of the given polynomials are

$$\mathbf{v}_1 := [p_1(x)]_B = \begin{bmatrix} 1 \\ 3 \\ 2 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_2 := [p_2(x)]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_3 := [p_3(x)]_B = \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_4 := [p_4(x)]_B = \begin{bmatrix} 3 \\ 8 \\ 0 \\ 8 \end{bmatrix}.$$

Let $T = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ be the set of these coordinate vectors.

---

We find a basis for $\operatorname{Span}(T)$ among the vectors in $T$ by the leading 1 method.

We form the matrix whose column vectors are the vectors in $T$ and apply elementary row operations as follows.

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 3 & 1 & 1 & 8 \\ 2 & 0 & 1 & 0 \\ -1 & 1 & -1 & 8 \end{bmatrix} \xrightarrow{\substack{R_2-3R_1 \\ R_3-2R_1 \\ R_4+R_1}} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & -6 \\ 0 & 1 & -1 & 11 \end{bmatrix} \xrightarrow{R_4-R_2}$$

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & -2 & 12 \end{bmatrix} \xrightarrow{R_4+2R_2} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The first three columns of the reduced row echelon form matrix contain the leading 1's.
Thus, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\operatorname{Span}(T)$.
It follows that $Q := \{p_1(x), p_2(x), p_3(x)\}$ is a basis for $\operatorname{Span}(S)$.

(b) For each polynomial in $S$ that is not in $Q$, find the coordinate vector with respect to the basis $Q$.

Note that $p_4(x)$ is not in the basis $Q$.

The fourth column of the matrix in reduced row echelon form of part (a) gives the coefficients of the linear combination:

$$p_4(x) = 3p_1(x) + 5p_2(x) - 6p_3(x).$$

Thus, the coordinate vector of $p_4(x)$ with respect to the basis $Q$ is

$$[p_4(x)]_Q = \begin{bmatrix} 3 \\ 5 \\ -6 \end{bmatrix}.$$

## Comment.

This is one of the midterm 2 exam problems for Linear Algebra (Math 2568) in Autumn 2017.

In part (b), some students stopped after obtaining the linear combination
$p_4(x) = 3p_1(x) + 5p_2(x) - 6p_3(x)$.
You must read the problem carefully. You are asked to find the coordinate vector of $p_4(x)$ with respect to $Q$.

10 Let $\mathbf{P}_2$ be the vector space of polynomials of degree $2$ or less.
**(a)** Prove that the set $\{1, 1+x, (1+x)^2\}$ is a basis for $\mathbf{P}_2$.
**(b)** Write the polynomial $f(x) = 2 + 3x - x^2$ as a linear combination of the basis $\{1, 1+x, (1+x)^2\}$.

## Proof.

(a) Prove that the set $\{1, 1+x, (1+x)^2\}$ is a basis for $\mathbf{P}_2$.

Consider the standard basis $\mathrm{B} = \{1, x, x^2\}$ of $\mathbf{P}_2$. Using this basis, we can write the elements using coordinate vectors as

$$[1]_{\mathrm{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad [1+x]_{\mathrm{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad [(1+x)^2]_{\mathrm{B}} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

We find the coordinate vector by writing an element as a linear combination of the basis elements. For example, $(1+x)^2 = 1 + 2x + 1x^2$, and so the coefficients $1, 2, 1$ translate into the column vector $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

Because $\dim \mathbf{P}_2 = 3$, this set is a basis if and only if these three vectors are linearly independent. To verify this, consider the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

This matrix is upper-triangular, and the diagonal entries are all non-zero. This implies the matrix is non-singular, and so the columns are linearly independent.

Thus, the set $\{1, 1+x, (1+x)^2\}$ is a basis of $\mathbf{P}_2$.

> (b) Write the polynomial $f(x) = 2 + 3x - x^2$ as a linear combination of the basis $\{1, 1+x, (1+x)^2\}$.

First Method: Using the technique of completing the square, we can factor the polynomial $f(x)$ as we like. Specifically,

$$\begin{aligned} f(x) &= -x^2 + 3x + 2 \\ &= -(x+1)^2 + 5x + 3 \\ &= -(x+1)^2 + 5(x+1) - 2. \end{aligned}$$

Hence, we have the linear combination

$$f(x) = -2 \cdot 1 + 5(1+x) - (1+x)^2. \tag{*}$$

---

Second Method: We can find this factorization by calculating the Taylor polynomial of $f(x)$ centered at $-1$.

This Taylor polynomial is defined by

$$f(x) = f(-1) + f'(-1)(x+1) + \frac{f''(-1)}{2}(x+1)^2.$$

The polynomial in (*) is recovered by finding $f(-1) = -2$, $f'(-1) = 5$, and $f''(-1) = -2$.

**11** Let $V$ be the vector space of all $2 \times 2$ real matrices. Let $S = \{A_1, A_2, A_3, A_4\}$, where

$$A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 4 \end{bmatrix}, A_3 = \begin{bmatrix} -1 & 0 \\ 1 & -10 \end{bmatrix}, A_4 = \begin{bmatrix} 3 & 7 \\ -2 & 6 \end{bmatrix}.$$

Then find a basis for the span $\mathrm{Span}(S)$.

### Solution.

Let $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ be the standard basis of $V$, where

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

We find the coordinate vectors of $A_1, A_2, A_3, A_4$ with respect to the basis $B$.

For the matrix $A_1$, we can write it as a linear combination of basis vectors of $B$ as follows.
We have

$$A_1 = 1 \cdot E_{11} + 2E_{12} + (-1)E_{21} + 3E_{22}.$$

Hence the coordinate vector of $A_1$ with respect to the basis $B$ is

$$[A_1]_B = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}.$$

Similarly, we obtain

$$[A_2]_B = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 4 \end{bmatrix}, [A_3]_B = \begin{bmatrix} -1 \\ 0 \\ 1 \\ -10 \end{bmatrix}, [A_4]_B = \begin{bmatrix} 3 \\ 7 \\ -2 \\ 6 \end{bmatrix}.$$

---

Let

$$\begin{aligned} T &= \{[A_1]_B, [A_2]_B, [A_3]_B, [A_4]_B\} \\ &= \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ -10 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ -2 \\ 6 \end{bmatrix} \right\}. \end{aligned}$$

Then $T$ is a subset in $\mathbb{R}^4$.

---

We find a basis of $\operatorname{Span}(T)$ by the leading $1$-method.
We form a matrix whose columns are vectors in $T$ and reduce it by elementary row operations as follows.

$$\begin{aligned} &\begin{bmatrix} 1 & 0 & -1 & 3 \\ 2 & -1 & 0 & 7 \\ -1 & 1 & 1 & -2 \\ 3 & 4 & -10 & 6 \end{bmatrix} \xrightarrow{\substack{R_2-2R_1 \\ R_3+R_1 \\ R_4-3R_1}} \begin{bmatrix} 1 & 0 & -1 & 3 \\ 0 & -1 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 4 & -7 & -3 \end{bmatrix} \\ &\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 2 & 1 \\ 0 & 4 & -7 & -3 \end{bmatrix} \xrightarrow{\substack{R_3+R_2 \\ R_4-4R_2}} \begin{bmatrix} 1 & 0 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & -7 & -7 \end{bmatrix} \\ &\xrightarrow{\substack{\frac{1}{2}R_3 \\ -\frac{1}{7}R_4}} \begin{bmatrix} 1 & 0 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{R_1+R_3 \\ R_4-R_3}} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

The last matrix is in reduced row echelon form and the first three columns contain the leading $1$.

Thus the first three vectors of $T$ is a basis of $\operatorname{Span}(T)$.

Hence

$$\{[A_1]_B, [A_2]_B, [A_3]_B\}$$

is a basis of $\operatorname{Span}(T)$.

This yields that

$$\{A_1, A_2, A_3\}$$

is a basis of $\operatorname{Span}(S)$ by the correspondence of coordinate vectors.

**12** Let $C[-1,1]$ be the vector space over $\mathbb{R}$ of all continuous functions defined on the interval $[-1,1]$. Let

$$V := \{f(x) \in C[-1,1] \mid f(x) = ae^x + be^{2x} + ce^{3x}, a, b, c \in \mathbb{R}\}$$

be a subset in $C[-1,1]$.
**(a)** Prove that $V$ is a subspace of $C[-1,1]$.
**(b)** Prove that the set $B = \{e^x, e^{2x}, e^{3x}\}$ is a basis of $V$.
**(c)** Prove that $B' = \{e^x - 2e^{3x}, e^x + e^{2x} + 2e^{3x}, 3e^{2x} + e^{3x}\}$ is a basis for $V$.

## Proof.

### (a) Prove that $V$ is a subspace of $C[-1,1]$.

Note that each function in the subset $V$ is a linear combination of the functions $e^x, e^{2x}, e^{3x}$.

Namely, we have

$$V = \operatorname{Span}\{e^x, e^{2x}, e^{3x}\}$$

and we know that the span is always a subspace. Hence $V$ is a subspace of $C[-1,1]$.

### (b) Prove that the set $B = \{e^x, e^{2x}, e^{3x}\}$ is a basis of $V$.

We noted in part (a) that $V = \operatorname{Span}(B)$. So it suffices to show that $B$ is linearly independent.

Consider the linear combination

$$c_1e^x + c_2e^{2x} + c_3e^{3x} = \theta(x),$$

where $\theta(x)$ is the zero function (the zero vector in $V$).

Taking the derivative, we get

$$c_1e^x + 2c_2e^{2x} + 3c_3e^{3x} = \theta(x).$$

Taking the derivative again, we obtain

$$c_1e^x + 4c_2e^{2x} + 9c_3e^{3x} = \theta(x).$$

Evaluating at $x = 0$, we obtain the system of linear equations

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ c_1 + 2c_2 + 3c_3 &= 0 \\ c_1 + 4c_2 + 9c_3 &= 0. \end{aligned}$$

---

We reduce the augmented matrix for this system as follows:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 4 & 9 & 0 \end{array}\right] \xrightarrow[R_3-R_1]{R_2-R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 3 & 8 & 0 \end{array}\right] \xrightarrow[R_3-3R_2]{R_1-R_2}$$
$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right] \xrightarrow{\frac{1}{2}R_3} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow[R_2-2R_2]{R_1+R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

It follows that the solution of the system is $c_1 = c_2 = c_3 = 0$.
Hence the set $B$ is linearly independent, and thus $B$ is a basis for $V$.

Anotehr approach.

---

Alternatively, we can show that the coefficient matrix is nonsingular by using the Vandermonde determinant formula as follows.
Observe that the coefficient matrix of the system is a Vandermonde matrix:

$$A := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1^2 & 2^2 & 3^2 \end{bmatrix}.$$

The Vandermonde determinant formula yields that

$$\det(A) = (3-1)(3-2)(2-1) = 2 \neq 0.$$

Hence the coefficient matrix $A$ is nonsingular.
Thus we obtain the solution $c_1 = c_2 = c_3 = 0$.

> (c) Prove that $B' = \{e^x - 2e^{3x}, e^x + e^{2x} + 2e^{3x}, 3e^{2x} + e^{3x}\}$ is a basis for $V$.

We consider the coordinate vectors of vectors in $B'$ with respect to the basis $B$.
The coordinate vectors with respect to basis $B$ are

$$[e^x - 2e^{3x}]_B = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, [e^x + e^{2x} + 2e^{3x}]_B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, [3e^{2x} + e^{3x}]_B = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}.$$

Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ be these vectors and let $T = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.
Then we know that $B'$ is a basis for $V$ if and only if $T$ is a basis for $\mathbb{R}^3$.

---

We claim that $T$ is linearly independent.
Consider the matrix whose column vectors are $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 3 \\ -2 & 2 & 1 \end{bmatrix} \xrightarrow{R_3+2R_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 4 & 1 \end{bmatrix} \xrightarrow[R_3-4R_1]{R_1-R_2}$$

$$\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & -11 \end{bmatrix} \xrightarrow{-\frac{1}{11}R_3} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_2-3R_3]{R_1+3R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

---

Thus, the matrix is nonsingular and hence the column vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.
As $T$ consists of three linearly independent vectors in the three-dimensional vector space $\mathbb{R}^3$, we conclude that $T$ is a basis for $\mathbb{R}^3$.
Therefore, by the correspondence of the coordinates, we see that $B'$ is a basis for $V$.

## Related Question.

If you know the Wronskian, then you may use the Wronskian to prove that the exponential functions $e^x, e^{2x}, e^{3x}$ are linearly independent.

See the post
Using the Wronskian for Exponential Functions, Determine Whether the Set is Linearly Independent for the details.

---

Try the next more general question.

**Problem**.
Let $c_1, c_2, \dots, c_n$ be mutually distinct real numbers.

Show that exponential functions

$$e^{c_1 x}, e^{c_2 x}, \dots, e^{c_n x}$$

are linearly independent over $\mathbb{R}$.

The solution is given in the post ↴
Exponential Functions are Linearly Independent

13 Let $P_n(\mathbb{R})$ be the vector space over $\mathbb{R}$ consisting of all degree $n$ or less real coefficient polynomials. Let

$$U = \{p(x) \in P_n(\mathbb{R}) \mid p(1) = 0\}$$

be a subspace of $P_n(\mathbb{R})$. Find a basis for $U$ and determine the dimension of $U$.

## Solution.

Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be an element in $U$.
Since $p(1) = 0$, we have $a_n + a_{n-1} + \cdots + a_1 + a_0 = 0$. Thus we have

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x - (a_n + a_{n-1} + \cdots + a_1) \\ &= a_n(x^n - 1) + a_{n-1}(x^{n-1} - 1) + \cdots a_1(x - 1). \end{aligned}$$

---

Let us define

$$q_n(x) = x^n - 1, q_{n-1}(x) = x^{n-1} - 1, \ldots, q_1(x) = x - 1 \in U.$$

Then the above computation shows that we have

$$p(x) = a_n q_n(x) + a_{n-1}q_{n-1}(x) + \cdots a_1 q_1(x).$$

In other words, any element $p(x) \in U$ is a linear combination of $q_1(x), \ldots, q_n(x)$.
Hence $B := \{q_1(x), \ldots, q_n(x)\}$ is a spanning set for $U$.

---

We show that $B$ is a linearly independent set.
Consider a linear combination

$$c_1 q_1(x) + \cdots + c_n q_n(x) = \theta(x) := 0x^n + 0x^{n-1} + \cdots + 0x + 0.$$

Then we have

$$c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x - (c_1 + c_2 + \cdots + c_n) = 0.$$

Comparing the coefficients of both sides, we obtain

$$c_1 = c_2 = \cdots = c_n = 0$$

and thus $q_1(x), \ldots, q_n(x)$ are linearly independent, and hence $B = \{q_1(x), \ldots, q_n(x)\}$ is a basis for the subspace $U$. Since the dimension is the number of vectors in a basis, the dimension of $U$ is $n$.

14 Let $V$ be a vector space over $\mathbb{R}$ and let $B$ be a basis of $V$. Let $S = \{v_1, v_2, v_3\}$ be a set of vectors in $V$. If the coordinate vectors of these vectors with respect to the basis $B$ is given as follows, then find the dimension of $V$ and the dimension of the span of $S$.

$$[v_1]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, [v_2]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, [v_3]_B = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

## Solution.

### Coordinate vectors

Let us first recall the definition of coordinate vectors.
Suppose that $V$ is a vector space over $\mathbb{R}$ and let $B = \{v_1, v_2, \dots, v_n\}$ be a basis of $V$ (hence the dimension of $V$ is $n$).

Then any element of $v$ of $V$ can be written as

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n,$$

where $c_1, c_2, \dots, c_n$ are scalars in $\mathbb{R}$.

This expression is unique and the coordinate vector of $v$ with respect to the basis $B$ is defined as

$$[v]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Thus the coordinate vector $[v]_B$ is just a usual $n$-dimensional vector.

## Main part of the solution

Note that in the current problem, the coordinate vectors are $4$-dimensional vectors.
This implies that the basis $B$ consists of four vectors. Hence the dimension of $V$ is $4$.

By the correspondence of the coordinate vectors, the dimension of $\operatorname{Span}(S)$ is the same as the dimension of $\operatorname{Span}(T)$, where

$$\begin{aligned} T &= \{[v_1]_B, [v_2]_B, [v_2]_B\} \\ &= \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}. \end{aligned}$$

To find the dimension of $\operatorname{Span}(T)$, we need to find a basis of $\operatorname{Span}(T)$.
One way to do this is to note that the third vector is the sum of the first two vectors. Also, it’s clear that the first two vectors are linearly independent.

Thus, the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

is a basis of $\operatorname{Span}(T)$, hence the dimension of $\operatorname{Span}(T)$ is $2$.
We conclude that the dimension of $\operatorname{Span}(S)$ is $2$ as well.
(We can also conclude that the set $\{v_1, v_2\}$ is a basis of $\operatorname{Span}(S)$.)

## Another way to find a basis of $\operatorname{Span}(T)$

Here is another way to find a basis of $\mathrm{Span}(T)$. We can use the leading 1 method.

We form the matrix whose column vectors are the vectors in $T$:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that this matrix is already in reduced row echelon form.

The first two columns contain the leading 1's. Thus the first two columns form a basis of $\mathrm{Span}(T)$ (this is the leading 1 method).

Remark: In general we apply elementary row operations to reduce the matrix into a matrix in reduced row echelon form.

⑮ Let $V$ be a vector space and $B$ be a basis for $V$.

Let $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5$ be vectors in $V$.

Suppose that $A$ is the matrix whose columns are the coordinate vectors of $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5$ with respect to the basis $B$. After applying the elementary row operations to $A$, we obtain the following matrix in reduced row echelon form $\begin{bmatrix} 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

**(a)** What is the dimension of $V$?

**(b)** What is the dimension of $\mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5\}$?

*(The Ohio State University)*

## Solution.

### (a) What is the dimension of $V$?

Suppose the dimension of $V$ is $n$.

This means that the basis $B$ consists of $n$ vectors.

Then the coordinate of $\mathbf{w}_1$ with respect to $B$ is an $n$-dimensional vector $[\mathbf{w}_1]_B \in \mathbb{R}^n$.

Thus, the number of rows in the matrix $A$ is $n$.

As the elementary row operations do not change the number of rows, we see from the given matrix that the $A$ has four rows.

Thus, the dimension of $V$ is $4$.

### (b) What is the dimension of $\mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5\}$?

Note that the dimension of $W := \mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5\}$ is the same as the dimension of

$$W' := \mathrm{Span}\{[\mathbf{w}_1]_B, [\mathbf{w}_2]_B, [\mathbf{w}_3]_B, [\mathbf{w}_4]_B, [\mathbf{w}_5]_B\}.$$

Since the column vectors of $A$ are these coordinate vectors $[\mathbf{w}_i]_B$ and its reduced row echelon contains the leading 1's in the first two columns, we conclude that $\{[\mathbf{w}_1]_B, [\mathbf{w}_2]_B\}$ is a basis for $W'$ by the leading 1 method.

It follows that $\{\mathbf{w}_1, \mathbf{w}_2\}$ is a basis for $W$, and its dimension is $2$.

⑯ Let $V$ be the set of all $n \times n$ diagonal matrices whose traces are zero. That is,

$$V := \left\{ A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \;\middle|\; \begin{array}{l} a_{11}, \dots, a_{nn} \in \mathbb{C}, \\ \mathrm{tr}(A) = 0 \end{array} \right\}$$

Let $E_{ij}$ denote the $n \times n$ matrix whose $(i, j)$-entry is $1$ and zero elsewhere.
**(a)** Show that $V$ is a subspace of the vector space $M_n$ over $\mathbb{C}$ of all $n \times n$ matrices. (You may assume without a proof that $M_n$ is a vector space.)
**(b)** Show that matrices

$$E_{11} - E_{22},\, E_{22} - E_{33},\, \dots,\, E_{n-1\,n-1} - E_{nn}$$

are a basis for the vector space $V$.
**(c)** Find the dimension of $V$.

## Hint (Subspace Criteria)

For (a), use the criteria for a subset to be a subspace.

⑰ The zero vector $\mathbf{0} \in M_n$ is in $V$.

⑱ For $\mathbf{u}, \mathbf{v} \in V$, the sum $\mathbf{u} + \mathbf{v} \in V$.

⑲ For $\mathbf{v} \in V, c \in \mathbb{C}$, the scalar product $c\mathbf{v} \in V$.

## Proof.

### (a) $V$ is a subspace of the vector space $M_n$ over $\mathbb{C}$

We check the following criteria for a subset to be a subspace.

⑳ The zero vector $\mathbf{0} \in M_n$ is in $V$.

㉑ For $\mathbf{u}, \mathbf{v} \in V$, the sum $\mathbf{u} + \mathbf{v} \in V$.

㉒ For $\mathbf{v} \in V, c \in \mathbb{C}$, the scalar product $c\mathbf{v} \in V$.

For (1), note that the zero vector $\mathbf{0} \in M_n$ is the $n \times n$ zero matrix. Thus it is clearly in $V$.

To check (2), take $\mathbf{u}, \mathbf{v} \in V$. Then $\mathbf{u}, \mathbf{v}$ are diagonal matrices and $\operatorname{tr}(\mathbf{u}) = 0, \operatorname{tr}(\mathbf{v}) = 0$.
Thus the sum $\mathbf{u} + \mathbf{v}$ is a diagonal matrix and the trace is

$$\operatorname{tr}(\mathbf{u} + \mathbf{v}) = \operatorname{tr}(\mathbf{u}) + \operatorname{tr}(\mathbf{v}) = 0 + 0 = 0.$$

Hence the sum $\mathbf{u} + \mathbf{v} \in V$.

Finally, to prove (3) we take $\mathbf{v} \in V, c \in \mathbb{C}$.
Then $\mathbf{v}$ is a diagonal matrix and $\operatorname{tr}(\mathbf{v}) = 0$.
Thus $c\mathbf{v}$ is also a diagonal matrix and the trace is

$$\operatorname{tr}(c\mathbf{v}) = c\operatorname{tr}(\mathbf{v}) = c \cdot 0 = 0.$$

Hence $c\mathbf{v} \in V$.
Therefore the subspace criteria (1)-(3) hold. Thus $V$ is a subspace of the vector space $M_n$.

### (b) A basis for the vector space $V$.

Note that for each $1 \leq i \leq n-1$, the matrix $E_{ii} - E_{i+1\, i+1}$ is a diagonal matrix and the trace is

$$\operatorname{tr}(E_{ii} - E_{i+1\, i+1}) = \operatorname{tr}(E_{ii}) - \operatorname{tr}(E_{i+1\, i+1}) = 1 - 1 = 0.$$

Hence the matrix $E_{ii} - E_{i+1\, i+1} \in V$.

We show that the matrices $E_{ii} - E_{i+1\, i+1}$ span $V$.
Let $A = (a_{ij}) \in V$. We want to solve the following linear combination.

$$A = c_1(E_{11} - E_{22}) + c_2(E_{22} - E_{33}) + \cdots + c_{n-1}(E_{n-1\, n-1} - E_{nn}).$$

Rearranging this, we obtain

$$A = c_1 E_{11} + (c_2 - c_1)E_{22} + \cdots + (c_{n-1} - c_{n-2})E_{n-1\, n-1} - c_{n-1}E_{nn}.$$

Namely, in the matrix form we have

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} c_1 & 0 & \dots & \dots & 0 \\ 0 & c_2 - c_1 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c_{n-1} - c_{n-2} & 0 \\ 0 & \dots & \dots & 0 & -c_{n-1} \end{bmatrix}.$$

Comparing entries, we have the system of $n$ equations in $n-1$ unknowns $c_1, \dots, c_{n-1}$

$$\begin{aligned} a_{11} &= c_1 \\ a_{22} &= c_2 - c_1 \\ a_{33} &= c_3 - c_2 \\ &\vdots \\ a_{n-1n-1} &= c_{n-1} - c_{n-2} \\ a_{nn} &= -c_{n-1}. \end{aligned}$$

---

We solve this system as follows.

The matrix form of this system is

$$\begin{bmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & & -1 & 1 \\ & & & & & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{nn} \end{bmatrix}.$$

(The empty entries are all zero.)

---

The augmented matrix is

$$\left[\begin{array}{cccccc|c} 1 & & & & & & a_{11} \\ -1 & 1 & & & & & a_{22} \\ & -1 & 1 & & & & a_{33} \\ & & \ddots & \ddots & & & \vdots \\ & & & & -1 & 1 & a_{n-1n-1} \\ & & & & & -1 & a_{nn} \end{array}\right].$$

Then we reduce this matrix as follows.

Add the first row to the second row, then add the second row to the third row.

Repeating this process we get

$$\left[\begin{array}{ccccc|r} 1 & & & & & a_{11} \\ & 1 & & & & a_{11} + a_{22} \\ & & 1 & & & a_{11} + a_{22} + a_{33} \\ & & \ddots & \ddots & & \vdots \\ & & & & 1 & a_{11} + a_{22} + \cdots + a_{n-1n-1} \\ & & & & 0 & a_{11} + a_{22} + \cdots + a_{nn} \end{array}\right].$$

---

Note that since $A \in V$, $\mathrm{tr}(A) = a_{11} + a_{22} + \cdots a_{nn} = 0$, thus the right-bottom entry ($(n, n)$-entry) of the augmented matrix is $0$, hence the system is consistent and it has the unique solution

$$\begin{aligned} c_1 &= a_{11} \\ c_2 &= a_{11} + a_{22} \\ c_3 &= a_{11} + a_{22} + a_{33} \\ &\vdots \\ c_{n-1} &= a_{11} + a_{22} + \cdots + a_{n-1n-1}. \end{aligned} \tag{*}$$

Therefore the matrices $E_{11} - E_{22}, E_{22} - E_{33}, \dots, E_{n-1\,n-1} - E_{nn}$ span $V$.

---

To show that these matrices are linearly independent, we can reuse the above computation.

Suppose we have a linear combination

$$\mathbf{0} = c_1(E_{11} - E_{22}) + c_2(E_{22} - E_{33}) + \cdots + c_{n-1}(E_{n-1\,n-1} - E_{nn}).$$

Here $\mathbf{0}$ is the $n \times n$ zero matrix.

Since $a_{ii} = 0$ in the above computation, we see from (*) that $c_1 = c_2 = \cdots = c_{n-1} = 0$, hence the matrices $E_{ii} - E_{i+1i+1}$ are linearly independent.

Therefore those matrices are a basis of $V$.

### (c) The dimension of $V$.

From part (b), the $n - 1$ matrices

$$E_{11} - E_{22}, E_{22} - E_{33}, \dots, E_{n-1\,n-1} - E_{nn}$$

are a basis of $V$. Thus the dimension of $V$ is $n - 1$.

23 Let $C[-2\pi, 2\pi]$ be the vector space of all real-valued continuous functions defined on the interval $[-2\pi, 2\pi]$.
Consider the subspace $W = \operatorname{Span}\{\sin^2(x), \cos^2(x)\}$ spanned by functions $\sin^2(x)$ and $\cos^2(x)$.
**(a)** Prove that the set $B = \{\sin^2(x), \cos^2(x)\}$ is a basis for $W$.
**(b)** Prove that the set $\{\sin^2(x) - \cos^2(x), 1\}$ is a basis for $W$.

## Solution.

### (a) Prove that the set $B = \{\sin^2(x), \cos^2(x)\}$ is a basis for $W$.

By definition of the subspace $W = \operatorname{Span}\{\sin^2(x), \cos^2(x)\}$, the we know that $B$ is a spanning set for $W$.

Thus, it remains to show that $B$ is linearly independent set.

Suppose that

$$c_1 \sin^2(x) + c_2 \cos^2(x) = 0.$$

This equality is true for all $x \in [-2\pi, 2\pi]$.

---

In particular, evaluating at $x=0$, we see that $c_2=0$.
Also, plugging in $x=\pi/2$ yields $c_1=0$.

Therefore, $\sin^2(x)$ and $\cos^2(x)$ are linearly independent, that is, $B$ is linearly independent.
As $B$ is a linearly independent spanning set, it is a basis for $W$.

### Prove that the set $\{\sin^2(x)-\cos^2(x), 1\}$ is a basis for $W$.

Note that $\sin^2(x)-\cos^2(x)$ and $1$ are both in $W$ since both functions are linear combination of $\sin^2(x)$ and $\cos^2(x)$. Here, we used the trigonometric identity $1=\sin^2(x)+\cos^{(} x)$.

By part (a), we see that $\dim(W)=2$. So if we show that the functions $\sin^2(x)-\cos^2(x)$ and $1$ are linearly independent, then they form a basis for $W$.

---

We consider the coordinate vectors of these functions with respect to the basis $B$.
We have

$$\begin{aligned}[\sin^2(x)-\cos^2(x)]_B &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and} \\ [1]_B = [\sin^2(x)+\cos^2(x)]_B &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}.\end{aligned}$$

Since we have

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \xrightarrow{R_2+R_1} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1-R_1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

the coordinate vectors are linearly independent, and hence $\sin^2(x)-\cos^2(x)$ and $1$ are linearly independent.
We conclude that $\{\sin^2(x)-\cos^2(x), 1\}$ is a basis for $W$.

(Another reasoning is that since the coordinate vectors form a basis for $\mathbb{R}^2$, $\{\sin^2(x)-\cos^2(x), 1\}$ is a basis for $W$.)

### Comment

You may directly show that $\{\sin^2(x)-\cos^2(x), 1\}$ is linearly independent just like we did for part (a).

㉔ Suppose that $B=\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^2$. Let $S:=[\mathbf{v}_1, \mathbf{v}_2]$. Note that as the column vectors of $S$ are linearly independent, the matrix $S$ is invertible. Prove that for each vector $\mathbf{v} \in V$, the vector $S^{-1}\mathbf{v}$ is the coordinate vector of $\mathbf{v}$ with respect to the basis $B$.

### Proof.

We first express the vector $\mathbf{v}$ as a linear combination of the basis vectors

$$\mathbf{v} = c_1\mathbf{v}_{+}c_2\mathbf{v}_2.$$

This expression is unique and the coordinate vector of $\mathbf{v}$ with respect to the basis $B$ is defined to be

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

---

Let

$$S^{-1}\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Or equivalently,

$$\mathbf{v} = S \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Our goal is to show that $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$.

---

We have

$$c_1\mathbf{v}_+c_2\mathbf{v}_2 = \mathbf{v} = S \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2.$$

Hence

$$(x_1 - c_1)\mathbf{v}_1 + (x_2 - c_2)\mathbf{v}_2 = \mathbf{0}.$$

As $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent, we obtain $x_1 = c_1$ and $x_2 = c_2$.

Therefore,

$$S^{-1}\mathbf{v} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = [\mathbf{v}]_B.$$

# Dimensions of General Vector Spaces

**Definition**

1. The **dimension** $\dim(V)$ of a vector space $V$ is the number of vectors in a basis for $V$.

**Summary**

Let $V$ be a vector space over a scalar field $K$. Suppose that \dim(V)=n. *Let*S=\{\mathbf{w}_1, \dots, \mathbf{w}_k\}*beasetofvectorsin*V$.

1. The dimension of $V$ does not depend on the choice of a basis.
2. If $W$ is a subspace of $V$, then $\dim(W) \leq \dim(V)$.
3. If $k > n$, then the set $S$ is linearly dependent.
4. If $k < n$, then the set $S$ cannot span the vector space $V$.

## Problems

1. Let $V$ be the vector space of all $2 \times 2$ matrices whose entries are real numbers. Let

$$W = \left\{ A \in V \;\middle|\; A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \text{ for any } a, b, c \in \mathbb{R} \right\}.$$

**(a)** Show that $W$ is a subspace of $V$.
**(b)** Find a basis of $W$.
**(c)** Find the dimension of $W$.
*(The Ohio State University)*

### Solution.

#### (a) Show that $W$ is a subspace of $V$.

To show that $W$ is a subspace of $V$, we verify the following subspace criteria.

2. The zero vector in $V$ is in $W$.

---

3. For all $A, B \in W$, the sum $A + B \in W$.

---

4. For all $A \in W$ and $r \in$ R, the scalar multiplication $rA \in W$.

Note that the zero vector in $V$ is the zero matrix $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

The zero matrix corresponds to the matrix in $W$ with $a = b = c = 0$, and hence the zero vector of $V$ is in $W$.

Condition 1 is met.

---

To verify condition 2, let

$$A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \text{ and } A' = \begin{bmatrix} a' & b' \\ c' & -a' \end{bmatrix}$$

be arbitrary elements in $W$, where $a, b, c, d, a', b', c', d' \in \mathrm{R}$.

Then we have

$$A + A' = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & -a' \end{bmatrix} = \begin{bmatrix} a + a' & b + b' \\ c + c' & -(a + a)' \end{bmatrix}$$

and this is of the form of elements of $W$. Hence $A + A' \in W$ and condition 2 is met.

---

Finally, let us check condition 3.

Let $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ be an arbitrary element in $W$ and let $r \in \mathrm{R}$ be any real number.

Then we have

$$rA = r \cdot \begin{bmatrix} a & b \\ c & -a \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & -(ra) \end{bmatrix}$$

and this is of the form of the elements of $W$. Thus $rA \in W$ and condition 3 is met.

Therefore, by the subspace criteria, we conclude that $W$ is a subspace of $V$.

### (b) Find a basis of $W$.

Any vector $A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ in the subspace $W$ can be written as

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

It follows that the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

span the subspace $W$.

If the above linear combination is the zero matrix, then $a = b = c = 0$.
This implies that these matrices are linearly independent.

Thus,

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

is a basis for $W$.

### (c) Find the dimension of $W$.

As the basis we obtained contains three vectors, the dimension of $W$ is 3.

5 Let $V$ be the vector space of all $2 \times 2$ matrices. Let $W$ be a subset of $V$ consisting of all $2 \times 2$ skew-symmetric matrices. (Recall that a matrix $A$ is skew-symmetric if $A^{\mathrm{T}} = -A$.)
**(a)** Prove that the subset $W$ is a subspace of $V$.
**(b)** Find the dimension of $W$.
(*The Ohio State University*)

## Proof.

### (a) Prove that the subset $W$ is a subspace of $V$

To prove that $W$ is a subspace of $V$, we check the following subspace criteria.

(i) The zero vector $\mathbf{0} \in V$ is in $W$.

(ii) For any vectors $\mathbf{u}, \mathbf{v} \in W$, the sum $\mathbf{u} + \mathbf{v}$ is in $W$.

(iii) For any vector $\mathbf{u} \in W$ and any scalar $c \in \mathrm{R}$, the scalar product $c\mathbf{u} \in W$.

The zero vector in $V$ is the $2 \times 2$ zero vector

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since have

$$O^{\mathrm{T}} = O = -O,$$

the zero vector $O$ is a skew-symmetric matrix. Thus $O$ is in $W$, and hence condition (i) is met.

For condition (ii), consider $A, B \in W$. This means that $A, B$ are skew symmetric matrices and thus we have

$$A^{\mathrm{T}} = -A, B^{\mathrm{T}} = -B.$$

To show that $A + B \in W$, we compute as follows.

$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}} \overset{(*)}{=} -A + (-B) = -(A+B).$$

Thus we have $(A+B)^{\mathrm{T}} = -(A+B)$, and it follows that the matrix $A + B$ is skew-symmetric.
Hence $A + B \in W$ and condition (ii) is also met.

---

To check condition (iii), consider $A \in W$ and $c \in \mathrm{R}$.
We want to show that $cA \in W$, that is, we want to show that $cA$ is a skew-symmetric matrix.

We see this by computing as follows.

$$\begin{aligned}(cA)^{\mathrm{T}} &= cA^{\mathrm{T}}\\ &= c(-A) \text{ since } A \text{ is skew-symmetric}\\ &= -(cA).\end{aligned}$$

Thus we obtained $(cA)^{\mathrm{T}} = -(cA)$, and thus $cA$ is a skew-symmetric matrix as required.
Hence $cA \in W$ and condition (iii) is satisfied.

We confirmed all three subspace criteria (i)-(iii), and thus conclude that $W$ is a subspace of $V$.

### (b) Find the dimension of $W$

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2 × 2 matrix.

If $A$ is a skew-symmetric matrix, namely we have $A^{\mathrm{T}} = -A$, we have

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} = -\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Comparing entries of the matrices, we obtain

$$\begin{aligned} a &= -a \\ b &= -c \\ d &= -d. \end{aligned}$$

It follows that $a = 0$, $d = 0$, and $c = -b$.

Thus any skew-symmetric matrix is of the form

$$A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} = b\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Therefore, the subspace $W$ is spanned by

$$B = \left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$

and since the set $B$ consists of only one vector, it is linearly independent, and thus $B$ is a basis. From this, we conclude that the dimension of $W$ is 1.

---

**6** Let $P_4$ be the vector space consisting of all polynomials of degree $4$ or less with real number coefficients. Let $W$ be the subspace of $P_2$ by

$$W = \{p(x) \in P_4 \mid p(1) + p(-1) = 0 \text{ and } p(2) + p(-2) = 0\}.$$

Find a basis of the subspace $W$ and determine the dimension of $W$.

### Proof.

Let $p(x) \in W$ and write

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$$

for some coefficients $a_0, a_1, a_2, a_3, a_4 \in \mathrm{R}$.

We first compute the values of $p(x)$ at $x = \pm 1$ and $x = \pm 2$.

We have

$$\begin{aligned} p(1) &= a_0 + a_1 + a_2 + a_3 + a_4 \\ p(1) &= a_0 - a_1 + a_2 - a_3 + a_4 \\ p(2) &= a_0 + 2a_1 + 4a_2 + 8a_3 + 16a_4 \\ p(-2) &= a_0 - 2a_1 + 4a_2 - 8a_3 + 16a_4. \end{aligned}$$

Since $p(x)$ is in $W$, it satisfies

$$p(1) + p(-1) = 0 \text{ and } p(2) + p(-2) = 0,$$

and thus we have

$$\begin{aligned} p(1) + p(-1) &= 2a_0 + 2a_2 + 2a_4 = 0 \\ p(2) + p(-1) &= 2a_0 + 8a_2 + 32a_4 = 0. \end{aligned}$$

Dividing them by 2, we have the system of linear equations in $a_0, a_2, a_4$.

$$\begin{aligned} a_0 + a_2 + a_4 &= 0 \\ a_0 + 4a_2 + 16a_4 &= 0. \end{aligned}$$

We reduce the augmented matrix of the system by elementary row operations as follows.

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 4 & 16 & 0 \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 15 & 0 \end{bmatrix} \xrightarrow{\frac{1}{3}R_2}$$
$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 5 & 0 \end{bmatrix} \xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 0 & -4 & 0 \\ 0 & 1 & 5 & 0 \end{bmatrix}.$$

Hence the solutions to the system is

$$\begin{aligned} a_0 &= 4a_4 \\ a_2 &= -5a_4 \end{aligned}$$

and $a_4$ is a free variable.

---

Substituting these relations into $p(x)$, we obtain

$$\begin{aligned} p(x) &= 4a_4 + a_1x - 5a_4x^2 + a_3x^3 + a_4x^4 \\ &= a_1x + a_3x^3 + a_4(4 - 5x^2 + x^4). \end{aligned}$$

Let

$$\begin{aligned} p_1(x) &= x \\ p_2(x) &= x^3 \\ p_3(x) &= 4 - 5x^2 + x^4. \end{aligned}$$

These are vectors in $W$.
(To see this, you may directly check the defining relations, or set $a_1 = a_3 = 0$ and $a_4 = 1$ to get $p_3(x)$. Similarly for $p_1(x)$ and $p_2(x)$.)

Then by the above computations, any vector $p(x)$ in $W$ is a linear combination

$$p(x) = a_1p_1(x) + a_3p_2(x) + a_4p_3(x).$$

Thus, $\{p_1(x), p_2(x), p_3(x)\}$ is a spanning set of $W$.

---

Also, the vectors $p_1(x), p_2(x), p_3(x)$ are linearly independent.
In fact, if we have $c_1p_1(x) + c_2p_2(x) + c_3p_3(x) = 0$, then we have

$$\begin{aligned} 0 &= c_1p_1(x) + c_2p_2(x) + c_3p_3(x) \\ &= c_1x + c_2x^3 + c_3(4 - 5x^2 + x^4) \\ &= 4c_3 + c_1x - 5c_3x^2 + c_2x^3 + c_3x^4. \end{aligned}$$

Since the coefficients must be zero, we have $c_1 = c_2 = c_3 = 0$.
This proves that $p_1(x), p_2(x), p_3(x)$ are linearly independent.

---

Therefore the set $\{p_1(x), p_2(x), p_3(x)\}$ is linearly independent spanning set of the subspace $W$, hence it is a basis for $W$.
Thus the dimension of the subspace $W$ is 3.

---

7 Let $p_1(x), p_2(x), p_3(x), p_4(x)$ be (real) polynomials of degree at most $3$. Which (if any) of the following two conditions is sufficient for the conclusion that these polynomials are linearly dependent?
**(a)** At $1$ each of the polynomials has the value $0$. Namely $p_i(1) = 0$ for $i = 1, 2, 3, 4$.

**(b)** At $0$ each of the polynomials has the value $1$. Namely $p_i(0)=1$ for $i=1,2,3,4$.
(*University of California, Berkeley*)

### Hint.

---

8 The set $P_3$ consisting of all polynomials of degree $3$ or less is a vector space of degree $4$.

---

9 Consider a subset of $P_3$ satisfying the condition (a) and prove that its dimension is less than or equal to $3$.

---

## Proof.

### (a) At $1$ each of the polynomials has the value $0$.

We show that the condition (a) is sufficient. Let us denote by $P_3$ the vector space consisting of all polynomials (with the variable $x$) of degree $3$ or less. Then the dimension of $P_3$ is $4$.

Consider

$$W := \{p(x) \in P_3 \mid p(1)=0\}.$$

Then this is a subspace of $P_3$. (Check it or see another solution below.)

Since, for example, the polynomial $q(x)=x \in P_3$ is not in $W$, the subspace $W$ is a proper subspace of $P_3$. Hence $\dim(W) < \dim(P_3)=4$.
(Actually, the dimension is $3$, see another solution below.)

Since the dimension of $W$ is less than or equal to $3$, any four vectors in $W$ must be linearly dependent. Thus $p_i$ are linearly dependent.

### Another Solution of (a)

This solution is simpler than the previous solution.

Consider a map $P_3 \to \mathbb{R}$ sending $p(x)$ to $p(1)$ (evaluating at $x=1$).
Then this is a linear transformation. The kernel of this linear transformation is

$$W := \{p(x) \in P_3 \mid p(1)=0\}.$$

(As it is the kernel, $W$ is a subspace of $P_3$.)
By *the rank-nullity theorem*, we see that $\dim(W)=3$. Hence any four vectors in $W$ are linearly dependent.

### (b) At $0$ each of the polynomials has the value $1$.

We show that the condition (b) is not sufficient.
In fact, the following four vectors satisfy the condition (b) but they are linearly independent:
$1, 1+x, 1+x+x^2, 1+x+x^2+x^3$ .

(Observe that the method of the proof of (a) does not work because

$$V = \{p(x) \in P_3 \mid p(0)=1\}$$

is not a subspace.)

## Comment.

The second proof of (a) is simpler and I think better than the first proof.

The reason I wrote the first proof is that it is more accessible for the first year linear algebra students.

- Let $V$ be the vector space of all $3 \times 3$ real matrices. Let $A$ be the matrix given below and we define

$$W = \{M \in V \mid AM = MA\}.$$

That is, $W$ consists of matrices that commute with $A$. Then $W$ is a subspace of $V$. Determine which matrices are in the subspace $W$ and find the dimension of $W$.

**(a)**

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix},$$

where $a, b, c$ are distinct real numbers.

**(b)**

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix},$$

where $a, b$ are distinct real numbers.

## Solution.

### (a) Diagonal matrix with distinct diagonal entries

Let us first determine when a matrix $M$ commutes with $A$.

Let

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

and suppose that $AM = MA$:

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}.$$

Computing matrix products, we obtain

$$\begin{bmatrix} aa_{11} & aa_{12} & aa_{13} \\ ba_{21} & ba_{22} & ba_{23} \\ ca_{31} & ca_{32} & ca_{33} \end{bmatrix} = \begin{bmatrix} a_{11}a & a_{12}b & a_{13}c \\ a_{21}a & a_{22}b & a_{23}c \\ a_{31}a & a_{32}b & a_{33}c \end{bmatrix}.$$

Compare the $(1, 2)$ entries and we have $aa_{12} = ba_{12}$.

Since $a \neq b$, we must have $a_{12} = 0$.

Similarly, comparing the off-diagonal entries and noting $a, b, c$ are distinct, we find that off diagonal entries $a_{ij}, i \neq j$ must be 0.

Thus, $M$ commutes with $A$ if and only if

$$M = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}.$$

Therefore, the subspace $W$ consists of all $3 \times 3$ diagonal matrices:

$$W = \{W \in V \mid W \text{ is diagonal}\}.$$

Then it is easy to see that the set $\{E_{11}, E_{22}, E_{33}\}$ is a basis for $W$, where $E_{ij}$ is the $3 \times 3$ matrix whose $(i, j)$-entry is 1 and the other entries are zero. Thus the dimension of $W$ is 3.

## (b) Diagonal matrix two diagonal entries are the same

Now consider the case

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}.$$

Let

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

and compute $AM = MA$ as in part (a) (or you just need to replace $b, c$ in (*) by $a, b$, respectively) and obtain

$$\begin{bmatrix} aa_{11} & aa_{12} & aa_{13} \\ aa_{21} & aa_{22} & aa_{23} \\ ba_{31} & ba_{32} & ba_{33} \end{bmatrix} = \begin{bmatrix} a_{11}a & a_{12}a & a_{13}b \\ a_{21}a & a_{22}a & a_{23}b \\ a_{31}a & a_{32}a & a_{33}b \end{bmatrix}.$$

Comparing entries and noting $a \neq b$, we have

$$a_{13} = 0, a_{23} = 0, a_{31} = 0, a_{32} = 0.$$

Thus, $M$ commutes with $A$ is and only if

$$M = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix},$$

and hence the subspace $W$ consists of such matrices.

From this, we see that the set

$$\{E_{11}, E_{12}, E_{21}, E_{22}, E_{33}\}$$

is a basis for $W$, and we conclude that the dimension of $W$ is 5.

- Let $U$ and $V$ be finite dimensional subspaces in a vector space over a scalar field $K$. Then prove that $\dim(U+V) \leq \dim(U) + \dim(V)$.

## Definition (The sum of subspaces).

Recall that the sum of subspaces $U$ and $V$ is

$$U + V = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in U, \mathbf{y} \in V\}.$$

The sum $U + V$ is a subspace.

(See the post "The sum of subspaces is a subspace of a vector space" for a proof.)

## Proof.

Let $n = \dim(U)$ and $m = \dim(V)$.

Let

$$B_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$$

be a basis of the vector space $U$ and let

$$B_2 = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$$

be a basis of the vector space $V$.

An arbitrary element of the vector space $U + W$ is of the form $\mathbf{x} + \mathbf{y}$, where $\mathbf{x} \in U$ and $\mathbf{y} \in V$.

---

Since $B_1$ is a basis of $U$, we can write

$$\mathbf{x} = r_1\mathbf{u}_1 + \cdots + r_n\mathbf{u}_n$$

for some scalars $r_1, \dots, r_n \in K$.

Also, since $B_2$ is a basis of $V$, we can write

$$\mathbf{y} = s_1\mathbf{v}_1 + \cdots + s_m\mathbf{v}_m$$

for some scalars $s_1, \dots, s_m \in K$.

---

Thus, we have

$$\mathbf{x} + \mathbf{y} = r_1\mathbf{u}_1 + \cdots + r_n\mathbf{u}_n + s_1\mathbf{v}_1 + \cdots + s_m\mathbf{v}_m,$$

and hence $\mathbf{x} + \mathbf{y}$ is in the span $S := \mathrm{Span}(\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m)$.

Thus we have $U + W \subset S$ and it yields that

$$\dim(U + W) \leq \dim(S) \leq n + m = \dim(U) + \dim(V).$$

This completes the proof.

## Related Question.

Let $A$ and $B$ be $m \times n$ matrices.

Prove that

$$\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B).$$

- Let $A$ and $B$ be $m \times n$ matrices. Prove that $\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$.

## Proof.

Let

$$A = [\mathbf{a}_1, \dots, \mathbf{a}_n] \text{ and } B = [\mathbf{b}_1, \dots, \mathbf{b}_n],$$

where $\mathbf{a}_i$ and $\mathbf{b}_i$ are column vectors of $A$ and $B$, respectively.

Then the rank of the matrix $A$ is the dimension of the column space of $A$.
That is, we have

$$\operatorname{rank}(A) = \dim(\operatorname{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)).$$

Similarly, we have

$$\operatorname{rank}(B) = \dim(\operatorname{Span}(\mathbf{b}_1, \dots, \mathbf{b}_n))$$

and

$$\operatorname{rank}(A+B) = \dim((\operatorname{Span}(\mathbf{a}_1 + \mathbf{b}_1, \dots, \mathbf{a}_n + \mathbf{b}_n))$$

since $A + B = [\mathbf{a}_1 + \mathbf{b}_1, \dots, \mathbf{a}_n + \mathbf{b}_n]$.

---

We claim that

$$\operatorname{Span}(\mathbf{a}_1 + \mathbf{b}_1, \dots, \mathbf{a}_n + \mathbf{b}_n) \subset \operatorname{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) + \operatorname{Span}(\mathbf{b}_1, \dots, \mathbf{b}_n).$$

Any vector $\mathbf{x} \in \operatorname{Span}(\mathbf{a}_1 + \mathbf{b}_1, \dots, \mathbf{a}_n + \mathbf{b}_n)$ can be written as

$$\mathbf{x} = r_1(\mathbf{a}_1 + \mathbf{b}_1) + \cdots + r_n(\mathbf{a}_n + \mathbf{b}_n)$$

for some scalars $r_1, \dots, r_n$.

---

Thus we have

$$\begin{aligned}\mathbf{x} &= r_1(\mathbf{a}_1 + \mathbf{b}_1) + \cdots + r_n(\mathbf{a}_n + \mathbf{b}_n) \\ &= (r_1\mathbf{a}_1 + \cdots + r_n\mathbf{a}_n) + (r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n) \\ &\in \operatorname{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) + \operatorname{Span}(\mathbf{b}_1, \dots, \mathbf{b}_n),\end{aligned}$$

and hence the claim is proved.

Then we have

$$\begin{aligned}&\operatorname{rank}(A+B)\\&=\dim(\,(\operatorname{Span}(\mathbf{a}_1+\mathbf{b}_1,\dots,\mathbf{a}_n+\mathbf{b}_n)\,)\\&\leq \dim(\,\operatorname{Span}(\mathbf{a}_1,\dots,\mathbf{a}_n)+\operatorname{Span}(\mathbf{b}_1,\dots,\mathbf{b}_n)\,)\\&\leq \dim(\,\operatorname{Span}(\mathbf{a}_1,\dots,\mathbf{a}_n)\,)+\dim(\,\operatorname{Span}(\mathbf{b}_1,\dots,\mathbf{b}_n)\,)\\&=\operatorname{rank}(A)+\operatorname{rank}(B).\end{aligned}$$

Here we used the fact that for two subspaces $U$ and $V$ of a vector space, we have

$$\dim(U+V)\leq \dim(U)+\dim(V).$$

For a proof of this fact, see the post “Dimension of the sum of two subspaces“.

# Linear Transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$

**Definition**

① A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** if $T$ satisfies the following two linearity conditions: For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

1. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$
2. $T(c\mathbf{x}) = cT(\mathbf{x})$

② The **nullspace** $\mathcal{N}(T)$ of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is

$$\mathcal{N}(T) = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}_m\}.$$

③ The **nullity** of $T$ is the dimension of $\mathcal{N}(T)$.

④ The **range** $\mathcal{R}(T)$ of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is

$$\mathcal{R}(T) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = T(\mathbf{x}) \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

⑤ The **rank** of $T$ is the dimension of $\mathcal{R}(T)$.

⑥ The **matrix representation** of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is an $m \times n$ matrix $A$ such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$.

**Summary**

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

① $T(\mathbf{0}_n) = \mathbf{0}_m$, where $\mathbf{0}_n$ and $\mathbf{0}_m$ are the zero vectors in $\mathbb{R}^n$ and $R^m$, respectively.

② The matrix representation $A$ of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is given by $A = [T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)]$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the standard basis for $\mathbb{R}^n$.

③ If $A$ is the matrix representaiton of a linear transformation $T$, then

1. $\mathcal{N}(T) = \mathcal{N}(A)$ and $\mathcal{R}(T) = \mathcal{R}(A)$.
2. The nullity of $T$ is the same as the nullity of $A$.
3. The rank of $T$ is the same as the rank of $A$.

## Problems

① Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that

$$T(\mathbf{e}_1) = \mathbf{u}_1 \text{ and } T(\mathbf{e}_2) = \mathbf{u}_2,$$

where $\{\mathbf{e}_1, \mathbf{e}_2\}$ is the standard unit vectors of $\mathbb{R}^2$ and

$$\mathbf{u}_1 = \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} \text{ and } \mathbf{u}_2 = \begin{bmatrix} 8 \\ 2 \\ 6 \end{bmatrix}.$$

Then find $T\left( \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right)$.

---

## (a) Solution 1. (Using the Gram-Schmidt process)

It is straightforward to check that the vectors $\mathrm{v}_1, \mathrm{v}_2$ are linearly independent, and hence the set $S$ is a basis of $\mathrm{Span}(S)$.
Since the dot (inner) product of $\mathrm{v}_1$ and $\mathrm{v}_2$ is

$$\mathrm{v}_1 \cdot \mathrm{v}_2 = 1 \neq 0,$$

$S$ is not an orthogonal basis. We apply the Gram-Schmidt process to generate an orthogonal basis from the basis $S$.

The Gram-Schmidt process for two vectors is as follows. We define vectors $\mathrm{u}_1, \mathrm{u}_2$ by the following formula. Then $B = \{\mathrm{u}_1, \mathrm{u}_2\}$ is an orthogonal basis of $\mathrm{Span}(S)$.

$$\begin{aligned} \mathrm{u}_1 &:= \mathrm{v}_1 \\ \mathrm{u}_2 &:= \mathrm{v}_2 - \frac{\mathrm{u}_1 \cdot \mathrm{v}_2}{\mathrm{u}_1 \cdot \mathrm{u}_1} \mathrm{u}_1. \end{aligned} \tag{*}$$

Since we have

$$\mathrm{u}_1 \cdot \mathrm{v}_2 = \mathrm{v}_1 \cdot \mathrm{v}_2 = 1 \text{ and } \mathrm{u}_1 \cdot \mathrm{u}_1 = \mathrm{v}_1 \cdot \mathrm{v}_1 = 2,$$

we compute

$$\begin{aligned} \mathrm{u}_2 &= \mathrm{v}_2 - \frac{1}{2} \mathrm{u}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix}. \end{aligned}$$

Therefore the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} \right\}$$

is an orthogonal basis of $\mathrm{Span}(S)$ .

Note that scaling by a nonzero scalar does not change the orthogonality, the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is also an orthogonal basis of $\mathrm{Span}(S)$ , just in case you prefer not to have a fraction.

### (a) Solution 2. (Using a pattern of the Gram-Schmidt process)

Here is another solution using a partial information of the Gram-Schmidt process.

As in Solution 1, the set $S$ is a (non-orthogonal) basis of $\mathrm{Span}(S)$ .

We want to apply the Gram-Schmidt process but suppose we only remember the pattern of the Gram-Schmidt process. Namely, we want to define orthogonal vectors $\mathrm{u}, \mathrm{u}_2$ by

$$\begin{aligned} \mathrm{u}_1 &:= \mathrm{v}_1 \\ \mathrm{u}_2 &:= \mathrm{v}_2 + a\mathrm{u}_1. \end{aligned}$$

Here $a$ is some number, which is given in the Gram-Schmidt process (*) but we don’t remember.

We can still determine the number $a$ as follows. Since $\mathrm{u}_1$ and $\mathrm{u}_2$ will be orthogonal, we have

$$\begin{aligned} 0 = \mathrm{u}_1 \cdot \mathrm{u}_2 &= \mathrm{u}_1 \cdot (\mathrm{v}_2 + a\mathrm{u}_1) \\ &= \mathrm{u}_1 \cdot \mathrm{v}_2 + a\mathrm{u}_1 \cdot \mathrm{u}_1 \\ &= 1 + 2a. \end{aligned}$$

Hence, we obtain $a = -1/2$.

Then we determine

$$\begin{aligned} \mathrm{u}_2 &= \mathrm{v}_2 - \frac{1}{2}\mathrm{u}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix}. \end{aligned}$$

(So we could complete the Gram-Schmidt process even though we didn’t remember the details.)

Hence the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} \right\}$$

is an orthogonal basis of $\operatorname{Span}(S)$ .

## (b) Solution.

We first express the vector $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ as the linear combination

$$\begin{bmatrix} 3 \\ -2 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3\mathrm{e}_1 - 2\mathrm{e}_2.$$

Then we compute

$$\begin{aligned} T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) &= T(3\mathrm{e}_1 - 2\mathrm{e}_2) \\ &= 3T(\mathrm{e}_1) - 2T(\mathrm{e}_2) && \text{by linearity of } T \\ &= 3\begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 8 \\ 2 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 15 \\ 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 16 \\ 4 \\ 12 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -6 \end{bmatrix}. \end{aligned}$$

Therefore we have found

$$T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -1 \\ -6 \end{bmatrix}.$$

② Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $\mathbf{0}_n$ and $\mathbf{0}_m$ be zero vectors of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Show that $T(\mathbf{0}_n) = \mathbf{0}_m$.

## Steps.

③ Express $0_n = 0_n + 0_n$ or $0_n = 0 \cdot 0_n$ .

④ Calculate $T(0_n)$ using step 1 and the definition of linear transformation.

I will give three proofs.

## Proof 1.

Since $0_n = 0_n + 0_n$, we have

$$T(0_n) = T(0_n + 0_n) = T(0_n) + T(0_n),$$

where the second equality follows since $T$ is a linear transformation.

Subtracting $T(0_n)$ from both sides of the equality, we obtain $0_m = T(0_n)$.
Note that $0_m = T(0_n) - T(0_n)$ since $T(0_n)$ is a vector in $\mathbb{R}^m$.

### Proof 2.

Observe that we have $0 \cdot 0_n = 0_n$. (This is a scalar multiplication of the scalar $0$ and the vector $0_n$.

Now we have

$$T(0_n) = T(0 \cdot 0_n) = 0 \cdot T(0_n) = 0_m.$$

Here we used the one of the properties of the linear transformation $T$ in the second equality.

### Proof 3.

Note that $0_n = 0_n - 0_n$.
Thus we have

$$T(0_n) = T(0_n - 0_n) = T(0_n) - T(0_n) = 0_m,$$

where we used the linearity of $T$ in the second equality.
In the last equality, note that the vector $T(0_n)$ is $m$-dimensional vector.

⑤ Determine whether the function $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x+1 \\ 3y \end{bmatrix}$$

is a linear transformation.

### Solution.

The function $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a not a linear transformation.

Recall that every linear transformation must map the zero vector to the zero vector.

However, we have

$$T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 0+1 \\ 3 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

So the function $T$ does not map the zero vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to the zero vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Thus, $T$ is not a linear transformation.

### Another solution

Another way to see this is, for example, as follows.

Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

(In fact, you may take any two vectors.)

Then we have

$$T(\mathbf{u})+T(\mathbf{v})=T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)+T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}1\\2\\0\end{bmatrix}+\begin{bmatrix}1\\1\\3\end{bmatrix}=\begin{bmatrix}2\\3\\3\end{bmatrix}.$$

On the other hand, we have

$$T\left(\mathbf{u}+\mathbf{v}\right)=T\left(\begin{bmatrix}1\\1\end{bmatrix}\right)=\begin{bmatrix}2\\2\\3\end{bmatrix}.$$

Therefore, we see that

$$T(\mathbf{u})+T(\mathbf{v})\neq T\left(\mathbf{u}+\mathbf{v}\right),$$

and hence $T$ is not a linear transformation.

---

6 Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be a linear transformation.
Let

$$\mathbf{u}=\begin{bmatrix}1\\2\end{bmatrix},\mathbf{v}=\begin{bmatrix}3\\5\end{bmatrix}$$

be 2-dimensional vectors. Suppose that

$$\begin{aligned}T(\mathbf{u})&=T\left(\begin{bmatrix}1\\2\end{bmatrix}\right)=\begin{bmatrix}-3\\5\end{bmatrix},\\ T(\mathbf{v})&=T\left(\begin{bmatrix}3\\5\end{bmatrix}\right)=\begin{bmatrix}7\\1\end{bmatrix}.\end{aligned}$$

Let $\mathbf{w}=\begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2$. Find the formula for $T(\mathbf{w})$ in terms of $x$ and $y$.

---

## Solution.

Note that the vectors $\mathrm{u},\mathrm{v}$ are basis vectors for $\mathbb{R}^2$.

Thus we can write the vector $\mathrm{w}$ as a linear combination of $\mathrm{u}$ and $\mathrm{v}$.

Let

$$a_1\mathrm{u}+a_2\mathrm{v}=\mathrm{w}.$$

We want to determine $a_1$ and $a_2$.

So we consider the augmented matrix

$$\left[\begin{array}{cc|c}1&3&x\\2&5&y\end{array}\right].$$

Applying the elementary row operations, we obtain a reduced row echelon form matrix for this matrix as follows. (This is the Gauss-Jordan elimination method.)

$$\left[\begin{array}{cc|c} 1 & 3 & x \\ 2 & 5 & y \end{array}\right] \xrightarrow{R_2-2R_1} \left[\begin{array}{cc|c} 1 & 3 & x \\ 0 & -1 & y-2x \end{array}\right]$$
$$\xrightarrow{-R_2} \left[\begin{array}{cc|c} 1 & 3 & x \\ 0 & 1 & 2x-y \end{array}\right] \xrightarrow{R_1-3R_2} \left[\begin{array}{cc|c} 1 & 0 & x-3(2x-y) \\ 0 & 1 & 2x-y \end{array}\right].$$

Therefore we have

$$\begin{aligned} a_1 &= x-3(2x-y) = -5x+3y \\ a_2 &= 2x-y \end{aligned}$$

and the linear combination is

$$\mathrm{w} = (-5x+3y)\mathrm{u} + (2x-y)\mathrm{v}.$$

Now we use the linearity of the linear transformation $T$, we calculate $T(\mathrm{w})$ as follows.

$$\begin{aligned} T(\mathrm{w}) &= T((-5x+3y)\mathrm{u} + (2x-y)\mathrm{v}) \\ &= (-5x+3y)T(\mathrm{u}) + (2x-y)T(\mathrm{v}) \\ &= (-5x+3y)\begin{bmatrix} -3 \\ 5 \end{bmatrix} + (2x-y)\begin{bmatrix} 7 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 15x-9y+14x-7y \\ -25x+15y+2x-y \end{bmatrix} \\ &= \begin{bmatrix} 29x-16y \\ -23x+14y \end{bmatrix}. \end{aligned}$$

Thus the formula is

$$T(\mathrm{w}) = \begin{bmatrix} 29x-16y \\ -23x+14y \end{bmatrix}.$$

## Comment.

After you obtained the formula, check the formula with $\mathrm{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and see if you get $\begin{bmatrix} -3 \\ 5 \end{bmatrix}$.

If the formula gives a different vector, then you made some mistake.

To be more confident, do the same with $\mathrm{w} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

⑦ Let $\{\mathbf{v}_1, \mathbf{v}_2\}$ be a basis of the vector space $\mathbb{R}^2$, where $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. The action of a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ on the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ is given by

$$T(\mathbf{v}_1) = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \text{ and } T(\mathbf{v}_2) = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}.$$

Find the formula of $T(\mathbf{x})$, where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$.

## Solution 1 (linear combination)

Since we know the values of $T$ on the basis vectors $\mathbf{v}_1, \mathbf{v}_2$, if we express the vector $\mathbf{x}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, we can find $F(\mathbf{x})$ by the linearity of the linear transformation $T$.

So let us find the scalars $c_1, c_2$ such that

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2.$$

We write this as

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

The matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is invertible (as its determinant is $-2$) and its inverse matrix is

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Thus, we have

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} x+y \\ x-y \end{bmatrix} \end{aligned}$$

Therefore, we obtain the linear combination

$$\mathbf{x} = \frac{1}{2}(x+y)\mathbf{v}_1 + \frac{1}{2}(x-y)\mathbf{v}_2.$$

Now we compute $T(\mathbf{x})$ as follows.
We have

$$\begin{aligned} T(\mathbf{x}) &= T\left(\frac{1}{2}(x+y)\mathbf{v}_1 + \frac{1}{2}(x-y)\mathbf{v}_2\right) \\ &= \frac{1}{2}(x+y)T(\mathbf{v}_1) + \frac{1}{2}(x-y)T(\mathbf{v}_2) \qquad \text{by linearity of } T \\ &= \frac{1}{2}(x+y)\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} + \frac{1}{2}(x-y)\begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix} \\ &= \begin{bmatrix} x+y \\ 6x-2y \\ 8x-2y \end{bmatrix}. \end{aligned}$$

Hence the formula is

$$T(\mathrm{x}) = \begin{bmatrix} x+y \\ 6x-2y \\ 8x-2y \end{bmatrix}.$$

Solution 2 (Matrix representation)

In the second solution, we use the matrix representation for the linear transformation $T$.

Let $A$ be the matrix of $T$ with respect to the standard basis $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ of $\mathbb{R}^2$.

Thus, we have $T(\mathrm{x}) = A\mathrm{x}$ by definition.

To find the matrix $A$, we compute

$$\begin{aligned} A \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} &= A \begin{bmatrix} \mathrm{v}_1 & \mathrm{v}_2 \end{bmatrix} \\ &= \begin{bmatrix} A\mathrm{v}_1 & A\mathrm{v}_2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 4 & 8 \\ 6 & 10 \end{bmatrix} \end{aligned}$$

It follows that we have

$$\begin{aligned} A &= \begin{bmatrix} 2 & 0 \\ 4 & 8 \\ 6 & 10 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2 & 0 \\ 4 & 8 \\ 6 & 10 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 6 & -2 \\ 8 & -2 \end{bmatrix}. \end{aligned}$$

We now be able to find $T(\mathrm{x})$ as follows.

We have

$$\begin{aligned} T(\mathrm{x}) &= A\mathrm{x} \\ &= \begin{bmatrix} 1 & 1 \\ 6 & -2 \\ 8 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x+y \\ 6x-2y \\ 8x-2y \end{bmatrix}, \end{aligned}$$

which is, of course, the same formula that we obtained in solution 1.

⑧ Let $T:\mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation such that

$$T(\mathbf{e}_1)=\begin{bmatrix}1\\4\end{bmatrix}, T(\mathbf{e}_2)=\begin{bmatrix}2\\5\end{bmatrix}, T(\mathbf{e}_3)=\begin{bmatrix}3\\6\end{bmatrix},$$

where

$$\mathbf{e}_1=\begin{bmatrix}1\\0\\0\end{bmatrix}, \mathbf{e}_2=\begin{bmatrix}0\\1\\0\end{bmatrix}, \mathbf{e}_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

are the standard unit basis vectors of $\mathbb{R}^3$.

For any vector $\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\in\mathbb{R}^3$, find a formula for $T(\mathbf{x})$.

## Hint.

Recall the definition of a linear transformation $T:\mathbb{R}^3\to\mathbb{R}^2$. A map $T$ is a linear transformation if the map $T$ satisfies:

⑨ $T(\mathrm{u}+\mathrm{v})=T(\mathrm{u})+T(\mathrm{v})$ for any $\mathrm{u},\mathrm{v}\in\mathbb{R}^3$, and

⑩ $T(c\mathrm{v})=cT(\mathrm{v})$ for any $\mathrm{v}\in\mathbb{R}^3$ and $c\in\mathbb{R}$.

## Solution.

Using the standard unit basis vectors, any vector $\mathrm{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\in\mathbb{R}^3$ can be expressed as a linear combination

$$\mathrm{x}=x_1\mathrm{e}_1+x_2\mathrm{e}_2+x_3\mathrm{e}_3.$$

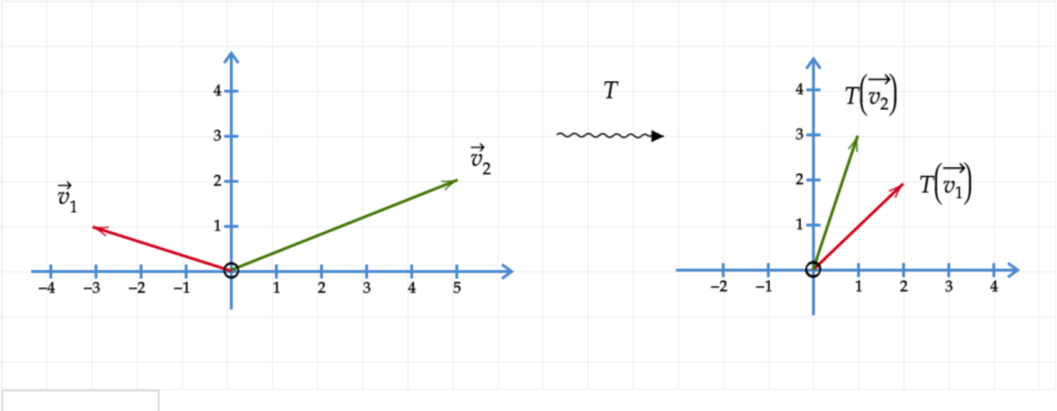
Since $T$ is a linear transformation, we have

$$\begin{aligned}T(\mathrm{x})&=T(x_1\mathrm{e}_1+x_2\mathrm{e}_2+x_3\mathrm{e}_3)\\&=x_1T(\mathrm{e}_1)+x_2T(\mathrm{e}_2)+x_3T(\mathrm{e}_3)\\&=x_1\begin{bmatrix}1\\4\end{bmatrix}+x_2\begin{bmatrix}2\\5\end{bmatrix}+x_3\begin{bmatrix}3\\6\end{bmatrix}\\&=\begin{bmatrix}x_1+2x_2+3x_3\\4x_1+5x_2+6x_3\end{bmatrix}.\end{aligned}$$

Therefore the formula is given by

$$T(\mathrm{x})=\begin{bmatrix}x_1+2x_2+3x_3\\4x_1+5x_2+6x_3\end{bmatrix}.$$

⑪ Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be a linear transformation such that it maps the vectors $\mathbf{v}_1,\mathbf{v}_2$ as indicated in the figure below.



Find the matrix representation $A$ of the linear transformation $T$.

## Solution 1.

From the figure, we see that

$$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix},$$

and

$$T(\mathbf{v}_1) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \text{ and } T(\mathbf{v}_2) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

---

Let $A$ be the matrix representation of the linear transformation $T$. By definition, we have $T(\mathbf{x}) = A\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^2$.

We determine $A$ as follows.

We have

$$\begin{aligned} \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} &= [T(\mathbf{v}_1), T(\mathbf{v}_2)] \\ &= [A\mathbf{v}_1, A\mathbf{v}_2] \\ &= A[\mathbf{v}_1, \mathbf{v}_2] \\ &= A \begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

---

Note that the determinant of the leftmost matrix is $-11$, hence it is invertible and the inverse is given by

$$\begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{11} \begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix}.$$

Hence

$$\begin{aligned} A &= \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix}^{-1} \\ &= \frac{1}{11} \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix}. \\ &= \frac{1}{11} \begin{bmatrix} -3 & 13 \\ -1 & 19 \end{bmatrix}. \end{aligned}$$

Therefore, the matrix representation of $T$ is

$$A = \frac{1}{11} \begin{bmatrix} -3 & 13 \\ -1 & 19 \end{bmatrix}.$$

## Solution 2.

Let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard basis for $\mathbb{R}^2$.
Then the matrix representation $A$ of the linear transformation $T$ is given by

$$A = [T(\mathbf{e}_1), T(\mathbf{e}_2)].$$

From the figure, we see that

$$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix},$$

and

$$T(\mathbf{v}_1) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \text{ and } T(\mathbf{v}_2) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

---

The values of $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ are not indicated in the figure but we can determine these values as follows.
Note that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^2$ (as they are linearly independent from the figure).

Let us express $\mathbf{e}_1$ as a linear combination of $\mathbf{v} - 1$ and $\mathbf{v}_2$.
Let

$$\mathbf{e}_1 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2,$$

where $c_1, c_2$ are scalars to be determined.

This is equivalent to

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = [\mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

As we have

$$\begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{11} \begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix},$$

it follows that

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Thus, we obtain the linear combination

$$\mathbf{e}_1 = -\frac{2}{11}\mathbf{v}_1 + \frac{1}{11}\mathbf{v}_2,$$

---

By the linearity of $T$, we have

$$\begin{aligned} T(\mathbf{e}_1) &= -\frac{2}{11}T(\mathbf{v}_1) + \frac{1}{11}T(\mathbf{v}_2) \\ &= -\frac{2}{11}\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \frac{1}{11}\begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ &= \frac{1}{11}\begin{bmatrix} -3 \\ -1 \end{bmatrix}. \end{aligned}$$

---

Similarly, let

$$\mathbf{e}_2 = d_1\mathbf{v}_1 + d_2\mathbf{v}_2,$$

where $d_1, d_2$ are scalars to be determined.
Then we have

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}.$$

It follows that

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} -3 & 5 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

Thus, we have the linear combination

$$\mathbf{e}_2 = \frac{5}{11}\mathbf{v}_1 + \frac{3}{11}\mathbf{v}_2,$$

and by linearity of $T$ we obtain

$$\begin{aligned} T(\mathbf{e}_2) &= \frac{5}{11}T(\mathbf{v}_1) + \frac{3}{11}T(\mathbf{v}_2) \\ &= \frac{5}{11}T(\mathbf{v}_1) + \frac{3}{11}T(\mathbf{v}_2) \\ &= \frac{5}{11}\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \frac{3}{11}\begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ &= \frac{1}{11}\begin{bmatrix} 13 \\ 19 \end{bmatrix}. \end{aligned}$$

---

In conclusion, the matrix representation for $T$ is

$$A=[T(\mathbf{e}_1),T(\mathbf{e}_2)]=\frac{1}{11}\begin{bmatrix}-3 & 13\\ -1 & 19\end{bmatrix}.$$

---

12 Let $T$ be a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ such that

$$T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right)=\begin{bmatrix}1\\2\end{bmatrix} \text{ and } T\left(\begin{bmatrix}0\\1\\1\end{bmatrix}\right)=\begin{bmatrix}0\\1\end{bmatrix}.$$

Then find $T\left(\begin{bmatrix}0\\1\\2\end{bmatrix}\right)$.

(*The Ohio State University*)

---

## Solution.

We first express the vector $\begin{bmatrix}0\\1\\2\end{bmatrix}$ as a linear combination

$$\begin{bmatrix}0\\1\\2\end{bmatrix}=c_1\begin{bmatrix}0\\1\\0\end{bmatrix}+c_2\begin{bmatrix}0\\1\\1\end{bmatrix}.$$

Then we find that $c_1=-1$ and $c_2=2$. Hence we obtain

$$\begin{bmatrix}0\\1\\2\end{bmatrix}=-\begin{bmatrix}0\\1\\0\end{bmatrix}+2\begin{bmatrix}0\\1\\1\end{bmatrix}.$$

---

We now compute

$$\begin{aligned}
T\left(\begin{bmatrix}0\\1\\2\end{bmatrix}\right)&=T\left(-\begin{bmatrix}0\\1\\0\end{bmatrix}+2\begin{bmatrix}0\\1\\1\end{bmatrix}\right)\\
&=-T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right)+2T\left(\begin{bmatrix}0\\1\\1\end{bmatrix}\right) && \text{by linearity of } T\\
&=-\begin{bmatrix}1\\2\end{bmatrix}+2\begin{bmatrix}0\\1\end{bmatrix}\\
&=\begin{bmatrix}-1\\0\end{bmatrix}.
\end{aligned}$$

Therefore we have found that

$$T\left(\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}-1\\0\end{bmatrix}$$

13 Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation such that

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, T(\mathbf{e}_2) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, T(\mathbf{e}_3) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the standard basis of $\mathbb{R}^3$.
Then find the rank and the nullity of $T$.
(*The Ohio State University*)

## Solution.

The matrix representation of the linear transformation $T$ is given by

$$A = [T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_3)] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that the rank and nullity of $T$ are the same as the rank and nullity of $A$.

The matrix $A$ is already in reduced row echelon form.

Thus, the rank of $A$ is $2$ because there are two nonzero rows.

---

Another way to see this is to use the leading 1 method. It implies that the first two columns vectors form a basis of the range of $A$ because the first two columns contain the leading 1's.

Thus, the rank (=the dimension of the range) is $2$.

---

The rank-nullity theorem says that

$$\text{rank of } A + \text{ nullity of } A = 3 \text{ (the number of columns of } A).$$

Hence the nullity of $A$ is $1$.

In summary, the rank of $T$ is $2$, and the nullity of $T$ is $1$.

14 Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that

$$T\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ and } T\left(\begin{bmatrix} 4 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}.$$

**(a)** Find the matrix representation of $T$ (with respect to the standard basis for $\mathbb{R}^2$).
**(b)** Determine the rank and nullity of $T$.
*(The Ohio State University)*

## Solution.

### (a) Find the matrix representation of $T$.

The matrix representation $A$ of the linear transformation is given by

$$A = [T(\mathrm{e}_1), T(\mathrm{e}_2)],$$

where

$$\mathrm{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathrm{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are the standard basis vectors for $\mathbb{R}^2$.

---

To determine $T(\mathrm{e}_1)$, we first express $\mathrm{e}_1$ as a linear combination of $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as follows.

Let

$$\mathrm{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = a \begin{bmatrix} 3 \\ 2 \end{bmatrix} + b \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

This can be written as

$$\begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The determinant of the coefficient matrix $\begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}$ is $3 \cdot 3 - 4 \cdot 2 = 1 \neq 0$ and thus it is invertible.

Hence we have

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

Hence $a = 3$ and $b = -2$.

It yields that

$$\mathrm{e}_1 = 3 \begin{bmatrix} 3 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

---

It follows that

$$\begin{aligned} T(\mathrm{e}_1) &= T\left(3 \begin{bmatrix} 3 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 3 \end{bmatrix}\right) \\ &= 3T\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right) - 2T\left(\begin{bmatrix} 4 \\ 3 \end{bmatrix}\right) && \text{by linearity of } T \\ &= 3 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 16 \\ 7 \end{bmatrix}. \end{aligned}$$

---

Similarly, we compute $T(\mathrm{e}_2)$ as follows.

Let

$$\mathrm{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = c \begin{bmatrix} 3 \\ 2 \end{bmatrix} + d \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

This can be written as

$$\begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Hence, we obtain

$$\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \end{bmatrix},$$

and $c = -4, d = 3$.

Hence

$$\mathrm{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -4\begin{bmatrix} 3 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

and we have

$$\begin{aligned} T(\mathrm{e}_2) &= T\left(-4\begin{bmatrix} 3 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 4 \\ 3 \end{bmatrix}\right) \\ &= -4T\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right) + 3T\left(\begin{bmatrix} 4 \\ 3 \end{bmatrix}\right) && \text{by linearity of } T \\ &= -4\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ -23 \\ -9 \end{bmatrix}. \end{aligned}$$

---

Therefore the matrix representation $A$ of $T$ is

$$A = [T(\mathrm{e}_1), T(\mathrm{e}_2)] = \begin{bmatrix} 3 & -4 \\ 16 & -23 \\ 7 & -9 \end{bmatrix}.$$

(b) Determine the rank and nullity of $T$.

Note that the rank and nullity of $T$ are the same as the those of the matrix representation $A$ of $T$.
In part (a), we obtained

$$A = \begin{bmatrix} 3 & -4 \\ 16 & -23 \\ 7 & -9 \end{bmatrix}.$$

Let us first determine the rank of $A$.
We have

$$A = \begin{bmatrix} 3 & -4 \\ 16 & -23 \\ 7 & -9 \end{bmatrix} \xrightarrow{R_3-2R_1} \begin{bmatrix} 3 & -4 \\ 16 & -23 \\ 1 & -1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & -1 \\ 16 & -23 \\ 3 & -4 \end{bmatrix}$$

$$\xrightarrow[R_3-3R_1]{R_2-16R_1} \begin{bmatrix} 1 & -1 \\ 0 & -7 \\ 0 & -1 \end{bmatrix} \xrightarrow{-R_3} \begin{bmatrix} 1 & -1 \\ 0 & -7 \\ 0 & 1 \end{bmatrix} \xrightarrow[R_2+7R_3]{R_1+R_3} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Hence, the reduced row echelon form matrix of $A$ has two nonzero rows.

So the rank of $A$ is $2$.

---

By the rank-nullity theorem, we know that

$$(\text{rank of } A) + (\text{nullity of } A) = 2.$$

As the rank of $A$ is $2$, we see that the nullity of $A$ is $0$.

15

Define the map $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \\ x_2 \end{bmatrix}$.

**(a)** Show that $T$ is a linear transformation.
**(b)** Find a matrix $A$ such that $T(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x} \in \mathbb{R}^2$.
**(c)** Describe the null space (kernel) and the range of $T$ and give the rank and the nullity of $T$.

Proof.

(a) Show that $T$ is a linear transformation.

To show that $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation, the map $T$ needs to satisfy:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, and
(ii) $T(c\mathbf{v}) = cT(\mathbf{v})$ for any $\mathbf{v} \in \mathbb{R}^2$ and $c \in \mathbb{R}$.

---

To check (i), let

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2.$$

We have

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) = T\left(\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}\right) &= \begin{bmatrix} (u_1 + v_1) - (u_2 + v_2) \\ (u_1 + v_1) + (u_2 + v_2) \\ u_2 + v_2 \end{bmatrix} \\ &= \begin{bmatrix} (u_1 - u_2) + (v_1 - v_2) \\ (u_1 + u_2) + (v_1 + v_2) \\ u_2 + v_2 \end{bmatrix} = \begin{bmatrix} u_1 - u_2 \\ u_1 + u_2 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 - v_2 \\ v_1 + v_2 \\ v_2 \end{bmatrix} \\ &= T(\mathbf{u}) + T(\mathbf{v}). \end{aligned}$$

Thus condition (i) holds.

---

To check (ii), consider any $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $c \in \mathbb{R}$.

Then we have

$$T(c\mathbf{v}) = T\left(\begin{bmatrix} cv_1 \\ cv_2 \end{bmatrix}\right) = \begin{bmatrix} cv_1 - cv_2 \\ cv_1 + cv_2 \\ cv_2 \end{bmatrix}$$

$$= c\begin{bmatrix} v_1 - v_2 \\ v_1 + v_2 \\ v_2 \end{bmatrix} = cT(\mathbf{v}).$$

Thus condition (ii) is met and the map $T$ is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$.

> (b) Find a matrix $A$ such that $T(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x} \in \mathbb{R}^2$.

**(b)** Let

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

be the standard basis vectors of $\mathbb{R}^2$.
The matrix $A$ satisfying $T(\mathbf{x}) = A\mathbf{x}$ can be obtained as

$$A = [T(\mathbf{e}_1) \quad T(\mathbf{e}_2)].$$

Since we have

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } T(\mathbf{e}_2) = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

we obtain the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

> (c) Describe the null space (kernel) and the range of $T$ and give the rank and the nullity of $T$

### Null Space and Nullity

We fist find the null space of the linear transformation of $T$.
Note that the null space of $T$ is the same as the null space of the matrix $A$.

By definition, the null space is

$$\mathcal{N}(T) = \mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^2 \mid A\mathbf{x} = \mathbf{0}\}.$$

So the null space is a set of all solutions for the system $A\mathbf{x} = \mathbf{0}$.

To find the solution of the system we consider the augmented matrix and reduce the matrix using the elementary row operations. (The Gauss-Jordan elimination method.)
We have

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right] \xrightarrow{R_2-R_1} \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[\text{then} R_2 \leftrightarrow R_3]{\substack{R_1+R_3 \\ R_2-2R_1}} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Thus the system has only zero solution

$$x_1 = 0, x_2 = 0.$$

Therefore we obtained

$$\mathcal{N}(T) = \mathcal{N}(A) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}.$$

Since the nullity is the dimension of the null space, we see that the nullity of $T$ is $0$ since the dimension of the zero vector space is $0$.

### Range and Rank

---

Next, we find the range of $T$. Note that the range of the linear transformation $T$ is the same as the range of the matrix $A$.
We describe the range by giving its basis.

The range of $A$ is the columns space of $A$. Thus it is spanned by columns

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

From the above reduction of the augmented matrix, we see that these vectors are linearly independent, thus a basis for the range. (Basically, this is the leading 1 method.)
Hence we have

$$\mathcal{R}(T) = \mathcal{R}(A) = \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\mathcal{R}(T)$.

The rank of $T$ is the dimension of the range $\mathcal{R}(T)$.
Thus the rank of $T$ is $2$.

Remark that we obtained that the nullity of $T$ is $0$ and the rank of $T$ is $2$. This agrees with the rank-nullity theorem

$$(\text{rank of } T) + (\text{nullity of } T) = 2.$$

---

**16** Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be a linear transformation defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 + 3x_3 - x_4 \\ 3x_1 + 5x_2 + 8x_3 - 2x_4 \\ x_1 + x_2 + 2x_3 \end{bmatrix}.$$

**(a)** Find a matrix $A$ such that $T(\mathbf{x}) = A\mathbf{x}$.
**(b)** Find a basis for the null space of $T$.
**(c)** Find the rank of the linear transformation $T$.
(The Ohio State University)

## Solution.

### (a) A matrix representation of a linear transformation

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, and $\mathbf{e}_4$ be the standard 4-dimensional unit basis vectors for $\mathrm{R}^4$.

Then the matrix representation for the linear transformation is given by the formula

$$A := \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) & T(\mathbf{e}_4) \end{bmatrix}.$$

Then we calculate

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}.$$

---

Similarly, we compute and obtain

$$T(\mathbf{e}_2) = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}, T(\mathbf{e}_3) = \begin{bmatrix} 3 \\ 8 \\ 2 \end{bmatrix}, T(\mathbf{e}_4) = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}.$$

Therefore the matrix $A$ we are looking for is

$$A = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 3 & 5 & 8 & -2 \\ 1 & 1 & 2 & 0 \end{bmatrix}.$$

### (b) The null space of the linear transformation

First, note that the null space of a linear transformation $T$ is the same as the null space of the matrix $A$ that represents $T$.

Thus, we find a basis for the null space of the matrix $A$ we obtained in (a).

The null space of $A$ consists of the solutions of $A\mathbf{x} = \mathbf{0}$ and so we apply Gauss-Jordan elimination process to solve the linear equation as follows. We apply the elementary row operations to the augmented matrix $[A \mid \mathbf{0}]$ of $A$.

$$[A \mid \mathbf{0}] = \begin{bmatrix} 1 & 2 & 3 & -1 & 0 \\ 3 & 5 & 8 & -2 & 0 \\ 1 & 1 & 2 & 0 & 0 \end{bmatrix} \xrightarrow[R_3-R_1]{R_2-3R_1} \begin{bmatrix} 1 & 2 & 3 & -1 & 0 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 \end{bmatrix}$$

$$\xrightarrow[\text{then } -R_2]{R_3-R_2} \begin{bmatrix} 1 & 2 & 3 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1-2R_2} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore the solutions are

$$\begin{aligned} x_1 &= -x_3 - x_4 \\ x_2 &= -x_3 + x_4 \end{aligned}$$

and thus the solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{0}$ is of the form

$$\mathbf{x} = \begin{bmatrix} -x_3 - x_4 \\ -x_3 + x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Hence we have

$$\mathcal{N}(T) = \mathcal{N}(A) = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

and the set

$$B := \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a spanning set for the null space $\mathcal{N}(T)$.

We claim that the set $B$ is linearly independent. Indeed, if we have a linear combination

$$a_1 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \mathbf{0},$$

then we have

$$\begin{bmatrix} -a_1 - a_2 \\ -a_1 + a_2 \\ a_1 \\ a_2 \end{bmatrix} = \mathbf{0}$$

and thus $a_1 = a_2 = 0$ and the set $B$ is linearly independent.

Since $B$ is a linearly independent spanning set for the null space $\mathcal{N}(T)$, the set $B$ is a basis for $\mathcal{N}(T)$.
In summary, we found a basis

$$B := \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

### (c) The rank of the linear transformation

We apply the rank-nullity theorem to find the rank of $T$.
From part (b), we see that nullity, the dimension of the null space $\mathcal{N}(T)$, is 2.

The linear transformation is a map from $R^4$ to $R^3$. Thus the rank-nullity theorem states that we have

$$\text{rank of } T + \text{nullity of } T = 4$$

$$\text{rank of } T + 2 = 4.$$

Therefore the rank of the linear transformation $T$ is 2.

---

17 Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that $T(\mathbf{e}_1) = \mathbf{u}_1$ and $T(\mathbf{e}_2) = \mathbf{u}_2$, where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are unit vectors of $\mathbb{R}^2$ and

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Then find $T\left( \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right)$.

### Hint.

A linear transformation from a vector space $V$ to a vector space $W$ is a map $f : V \to W$ satisfying the following linearity properties:

⑱ $f(u+v)=f(u)+f(v)$ for any vectors $u, v \in V$, and

⑲ $f(rv)=rf(v)$ for any vector $v \in V$ and any scalar $r$.

Note that the set $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for the vector space $\mathbb{R}^2$.

Thus the vector $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ can be written as a linear combination of the basis vectors $\mathbf{e}_1, \mathbf{e}_2$.

## Solution.

We first express the vector $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ as a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$:

$$\begin{bmatrix} 3 \\ -2 \end{bmatrix} = 3\mathbf{e}_1 - 2\mathbf{e}_2.$$

Then we have

$$\begin{aligned} T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) &= T(3\mathbf{e}_1 - 2\mathbf{e}_2) \\ &= 3T(\mathbf{e}_1) - 2T(\mathbf{e}_2) \text{ by the linearity of } T \\ &= 3\mathbf{u}_1 - 2\mathbf{u}_2 \\ &= 3\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -7 \\ -2 \\ 3 \end{bmatrix}. \end{aligned}$$

Thus, we found

$$T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) = \begin{bmatrix} -7 \\ -2 \\ 3 \end{bmatrix}.$$

⑳ Let $B=\{\mathbf{v}_1, \mathbf{v}_2\}$ be a basis for the vector space $\mathbb{R}^2$, and let $T:\mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that

$$T(\mathbf{v}_1) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \text{ and } T(\mathbf{v}_2) = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

If $\mathbf{e}_1 = \mathbf{v}_1 + 2\mathbf{v}_2$ and $\mathbf{e}_2 = 2\mathbf{v}_1 - \mathbf{u}_2$, where $\mathbf{e}_1, \mathbf{e}_2$ are the standard unit vectors in $\mathbb{R}^2$, then find the matrix of $T$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$.

## Solution.

The matrix representation $A$ of the linear transformation $T$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ is given by

$$A = [T(\mathbf{e}_1) \quad T(\mathbf{e}_2)]. \qquad (*)$$

To find the vectors $T(\mathbf{e}_1), T(\mathbf{e}_2)$ we use the linearity of $T$.

We have

$$\begin{aligned} T(\mathbf{e}_1) &= T(\mathbf{v}_1 + 2\mathbf{v}_2) \\ &= T(\mathbf{v}_1) + 2T(\mathbf{v}_2) \qquad \text{(by the linearity of } T\text{)} \\ &= \begin{bmatrix} 1 \\ -2 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 7 \\ 0 \end{bmatrix}. \end{aligned}$$

---

Also we have

$$\begin{aligned} T(\mathbf{e}_2) &= T(2\mathbf{v}_1 - \mathbf{u}_2) \\ &= 2T(\mathbf{v}_1) - T(\mathbf{u}_2) \\ &= 2\begin{bmatrix} 1 \\ -2 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ -5 \end{bmatrix}. \end{aligned}$$

---

Therefore the matrix $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix}$ of the linear transformation $T$ is given by

$$A = \begin{bmatrix} 7 & -1 \\ 0 & -5 \end{bmatrix}.$$

---

(21) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that

$$T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Then find the matrix $A$ such that $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^2$, and find the rank and nullity of $T$. (*The Ohio State University*)

## Solution.

The matrix $A$ such that $T(\mathbf{x}) = A\mathbf{x}$ is given by

$$A = [T(\mathbf{e}_1), T(\mathbf{e}_2)],$$

where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are standard basis of $\mathbb{R}^2$.

Since the vector $T(\mathbf{e}_2)$ is given, it remains to find $T(\mathbf{e}_1)$.

By inspection, we obtain the linear combination

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Thus, we have

$$\begin{aligned} T(\mathbf{e}_1) &= T\left(\begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}0\\1\end{bmatrix}\right) \\ &= T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) - T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) && \text{by linearity of } T \\ &= \begin{bmatrix}4\\1\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix} \\ &= \begin{bmatrix}1\\-1\end{bmatrix}. \end{aligned}$$

It follows that the matrix $A$ is given by

$$A = \begin{bmatrix}1 & 3\\-1 & 2\end{bmatrix}.$$

The rank and nullity of $T$ are the same as the rank and nullity of $A$.
We reduced the matrix $A$ by elementary row operations as follows:

$$A \xrightarrow{R_2+R_1} \begin{bmatrix}1 & 3\\0 & 5\end{bmatrix} \xrightarrow{\frac{1}{5}R_2} \begin{bmatrix}1 & 3\\0 & 1\end{bmatrix} \xrightarrow{R_1-3R_2} \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.$$

Hence the rank of $A$ is $2$ (because there are two non zero rows). The nullity of $A$ is determined by the rank nullity theorem

$$\text{rank of } A + \text{ nullity of } A = 2 \text{ (the number of columns of } A\text{)}.$$

Hence the nullity of $A$ is $0$.

In a nutshell, the rank of $T$ is $2$, and the nullity of $T$ is $0$.

22 If $L : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation such that

$$L\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\1\\2\end{bmatrix}, \quad L\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \begin{bmatrix}2\\3\\2\end{bmatrix}.$$

then

**(a)** find $L\left(\begin{bmatrix}1\\2\end{bmatrix}\right)$, and

**(b)** find the formula for $L\left(\begin{bmatrix}x\\y\end{bmatrix}\right)$.

(*Purdue University*)

## Hint.

23 Express $\begin{bmatrix}1\\2\end{bmatrix}$ as a linear combination of $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$.

㉔ Use the linearity of linear transformation $L$.

㉕ Same for part (b). Replace $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ with the general vector $\begin{bmatrix} x \\ y \end{bmatrix}$.

## Solution.

### (a) Find $L\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$

Note that the vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are a basis of $\mathbb{R}^2$.

We first express $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as a linear combination of the vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
Let $\begin{bmatrix} 1 \\ 2 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ c_2 \end{bmatrix}$.
Solving this, we have $c_1 = -1$ and $c_2 = 2$.

Then we calculate

$$\begin{aligned} L\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) &= L\left(-\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = -L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + 2L\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \\ &= -\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}, \end{aligned}$$

where the second equality follows from the linearity of $L$.

Thus we have

$$L\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}.$$

### (b) Find the formula for $L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$.

We generalize the proof of (a).
Let $\begin{bmatrix} x \\ y \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ c_2 \end{bmatrix}$ be a linear combination. Solving this, we have $c_1 = x - y, c_2 = y$.

Hence the linear combination is

$$\begin{bmatrix} x \\ y \end{bmatrix} = (x - y)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Using the linearity of $L$, we compute

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = L\left((x-y)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = (x-y)L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + yL\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

$$= (x-y)\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} x+y \\ x+2y \\ 2x \end{bmatrix}.$$

Therefore the formula is

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x+2y \\ 2x \end{bmatrix}.$$

**26** Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation satisfying

$$T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Find a general formula for $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$.

(*The Ohio State University*)

Solution 1. (Using linear combination)

Note that the set $B := \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ form a basis of the vector space $\mathbb{R}^2$.

---

To find a general formula, we first express the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ as a linear combination of the basis vectors in $B$.

Namely, we find scalars $c_1, c_2$ satisfying

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

This can be written as the matrix equation

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \tag{*}$$

where we have put

$$P = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

The $2 \times 2$ matrix $P$ is invertible as its determinant is $\det(P) = (1)(1) - (0)(2) = 1 \neq 0$, and the inverse matrix is given by

$$P^{-1} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}.$$

---

Then we multiply $P^{-1}$ on the left of (*) and obtain

$$\begin{aligned}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= P^{-1}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ -2x_1 + x_2 \end{bmatrix}.\end{aligned}$$

Thus, we have determined the scalars

$$c_1 = x_1 \text{ and } c_2 = -2x_1 + x_2,$$

and the linear combination becomes

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-2x_1 + x_2)\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

---

Now we compute $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$ as follows.

$$\begin{aligned}T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= T\left(x_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-2x_1 + x_2)\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= x_1 T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) + (-2x_1 + x_2)T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) && \text{by linearity of } T \\ &= x_1\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} + (-2x_1 + x_2)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 3x_1 \\ 4x_1 \\ 3x_1 + x_2 \end{bmatrix}.\end{aligned}$$

Therefore, the general formula is given by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 \\ 4x_1 \\ 3x_1 + x_2 \end{bmatrix}.$$

### Solution 2. (Using the matrix representation of the linear transformation)

The second solution uses the matrix representation of the linear transformation $T$.
Let $A$ be the matrix for the linear transformation $T$.

Then by definition, we have

$$T(\mathbf{x}) = A\mathbf{x}, \tag{**}$$

for every $\mathbf{x} \in \mathbb{R}^2$.
(Note that the size of $A$ is $3 \times 2$ because $T : \mathbb{R}^2 \to \mathbb{R}^3$.)

We determine the matrix $A$ as follows.

We compute

$$\begin{aligned} A\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} &= \begin{bmatrix} A\begin{bmatrix} 1 \\ 2 \end{bmatrix} & A\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) & T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \end{bmatrix} && \text{by (**)} \\ &= \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 5 & 1 \end{bmatrix}. \end{aligned}$$

The $2 \times 2$ matrix on the left hand side is invertible, and the inverse is

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}.$$

(Remark: This is the matrix $P$ in the first solution.)

Multiplying by this inverse matrix on the right, we obtain

$$\begin{aligned} A &= \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 3 & 1 \end{bmatrix}. \end{aligned}$$

---

Now that we have obtained the matrix $A$ for $T$, we can find the general formula as follows.

We have

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} && \text{by (**)} \\ &= \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} 3x_1 \\ 4x_1 \\ 3x_1 + x_2 \end{bmatrix}. \end{aligned}$$

---

In summary, the general formula is

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 \\ 4x_1 \\ 3x_1 + x_2 \end{bmatrix}.$$

---

27 Let $T$ be the linear transformation from the $3$-dimensional vector space $\mathbb{R}^3$ to $\mathbb{R}^3$ itself satisfying the following relations.

$$T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Then for any vector $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$, find the formula for $T(\mathbf{x})$.

Solution 1 using the matrix representation.

The first solution uses the matrix representation of $T$.

Let $A$ be the matrix representation of the linear transformation $T$ with respect to the standard basis of $\mathbb{R}^3$.

Then we have $T(\mathbf{x}) = A\mathbf{x}$ by definition.

We determine the matrix $A$ as follows.

We have

$$\begin{aligned} A\begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & 5 & 2 \end{bmatrix} &= \left[ A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ A\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \ A\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right] \\ &= \left[ T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ T\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \ T\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix}. \end{aligned}$$

It follows that we have

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & 5 & 2 \end{bmatrix}^{-1}.$$

Let us find the inverse matrix by using augmented matrix.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 1 & 5 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{R_2-R_1\\R_3-R_1}} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 3 & 2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{\substack{R_1-2R_2\\R_3-3R_2}} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 3 & -2 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 2 & -3 & 1 \end{array}\right] \xrightarrow{-R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 3 & -2 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array}\right]$$

$$\xrightarrow{\substack{R_1+2R_3\\R_2-R_3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 4 & -2 \\ 0 & 1 & 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array}\right].$$

Thus we obtain the inverse matrix

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & 5 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 4 & -2 \\ 1 & -2 & 1 \\ -2 & 3 & -1 \end{bmatrix},$$

and hence we have

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 4 & -2 \\ 1 & -2 & 1 \\ -2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} -3 & 7 & -3 \\ 2 & -4 & 2 \\ -2 & 6 & -3 \end{bmatrix}.$$

Using the relation $T(\mathbf{x}) = A\mathbf{x}$, the formula for $T(\mathbf{x})$ is obtained as follows.

$$\begin{aligned} T(\mathbf{x}) &= A\mathbf{x} \\ &= \begin{bmatrix} -3 & 7 & -3 \\ 2 & -4 & 2 \\ -2 & 6 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \begin{bmatrix} -3x + 7y - 3z \\ 2x - 4y + 2z \\ -2x + 6y - 3z \end{bmatrix}. \end{aligned}$$

### Solution 2 using a linear combination and linearity.

The second method is to find the linear combination

$$\mathbf{x} = c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

and use the linearity of the linear transformation.

To find the coefficients $c_1, c_2, c_3$, we consider the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & x \\ 1 & 3 & 1 & y \\ 1 & 5 & 2 & z \end{array}\right]$$

and we reduce this matrix by elementary row operations.

The reduction operations are exactly the same as in solution 1 and we obtain

$$\begin{aligned} c_1 &= -x + 4y - 2z \\ c_2 &= x - 2y + z \\ c_3 &= -2x + 3y - z. \end{aligned}$$

Therefore, we have by the linearity of $T$

$$\begin{aligned} T(\mathbf{x}) &= T\left( c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right) \\ &= c_1 T\left( \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right) + c_2 T\left( \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \right) + c_3 T\left( \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right) \\ &= (-x + 4y - 2z) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (x - 2y + z) \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} + (-2x + 3y - z) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -3x + 7y - 3z \\ 2x - 4y + 2z \\ -2x + 6y - 3z \end{bmatrix}. \end{aligned}$$

---

**28** Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation given by

$$T\left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} x_1 - x_2 \\ x_2 \\ x_1 + x_2 \end{bmatrix}.$$

Find an orthonormal basis of the range of $T$.
(*The Ohio State University*)

---

## Solution.

Let $A$ be the matrix representation of the linear transformation of $T$.

That is,

$$A = [\, T(\mathrm{e}_1) \quad T(\mathrm{e}_2) \,],$$

where

$$\mathrm{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathrm{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

form the standard basis of the vector space $R^2$.

By the formula, we see that

$$T\left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, T\left( \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

and thus the matrix $A$ for $T$ is

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Note that the range of $T$ is the same as the range of $A$.

We reduce the matrix $A$ by the elementary row operations as follows:

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \xrightarrow{R_3 - R_1} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \xrightarrow{R_1 + R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Since the both columns contain the leading $1$'s, we conclude that

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

is a basis of the range of $A$ by the leading $1$ method.

Note that the dot product of these basis vectors is $0$, hence they are already orthogonal.

Hence to obtain an orthonormal basis, we just need to normalize the length of these vectors to $1$.

In summary, an orthonormal basis of the range of $T$ is

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

This problem is one of the final exam problems of Linear Algebra course at the Ohio State University (Math 2568).

㉙ Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Suppose that $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is a subset of $\mathbb{R}^n$ such that $\{T(\mathbf{x}_1), T(\mathbf{x}_2), \dots, T(\mathbf{x}_k)\}$ is a linearly independent subset of $\mathbb{R}^m$. Prove that the set $S$ is linearly independent.

### Hint.

Recall the following definitions.

㉚ Vectors $\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_k$ are linearly independent
if and only if the only solution to the vector equation

$$c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_k\mathrm{x}_k = 0_n$$

is $c_1 = c_2 = \cdots = c_k = 0$.

㉛ A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a map such that

㉜ $T(\mathrm{v} + \mathrm{w}) = T(\mathrm{v}) + \mathrm{w}$ for all $\mathrm{v}, \mathrm{w} \in \mathbb{R}^n$.

(33) $T(c\mathbf{v}) = cT(\mathbf{v})$ for all $c \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^n$ .

## Proof.

Consider a linear combination of vectors in $S$

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0}_n,$$

where $\mathbf{0}_n$ is the $n$ dimensional zero vector.
To show that $S$ is linearly independent, we need to show that the coefficients $c_i$ are all zero.

Recall that any linear transformation maps the zero vector to the zero vector.
(See A linear transformation maps the zero vector to the zero vector for a proof of this fact.)

Thus we have

$$\begin{aligned} \mathbf{0}_m &= T(\mathbf{0}_n) = T(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) \\ &= c_1T(\mathbf{x}_1) + c_2T(\mathbf{x}_2) + \cdots + c_kT(\mathbf{x}_k). \end{aligned}$$

In the last step, we used the linearity of $T$ .

Since the vectors $T(\mathbf{x}_1), T(\mathbf{x}_2), \dots, T(\mathbf{x}_k)$ are linearly independent, the coefficient of this linear combination of these vectors must be zero.
Thus we have $c_1 = c_2 = \cdots = c_k = 0$, hence the set $S$ is linearly independent.

(34) Let $T$ be a linear transformation from the vector space $\mathbb{R}^3$ to $\mathbb{R}^3$.
Suppose that $k=3$ is the smallest positive integer such that $T^k = \mathbf{0}$ (the zero linear transformation) and suppose that we have $\mathbf{x} \in \mathbb{R}^3$ such that $T^2\mathbf{x} \neq \mathbf{0}$. Show that the vectors $\mathbf{x}, T\mathbf{x}, T^2\mathbf{x}$ form a basis for $\mathbb{R}^3$.
(*The Ohio State University*)

## Proof.

We first recall that $T(0) = 0$ . To see this, we compute

$$T(0) = T(0 \cdot 0) = 0 \cdot T(0) = 0.$$

To show that $\mathbf{x}, T\mathbf{x}, T^2\mathbf{x}$ form a basis, it suffices to prove that these vectors are linearly independent since the dimension of the vector space $\mathbb{R}^3$ is three and any three linearly independent vectors form a basis in such a vector space.
(For a different proof, see the post A linear transformation maps the zero vector to the zero vector.)

---

Let us consider a linear combination

$$a_1\mathbf{x} + a_2(T\mathbf{x}) + a_3(T^2\mathbf{x}) = 0, \tag{*}$$

where $a_1, a_2, a_3$ are real scalars.

If the only solution is $a_1 = a_2 = a_3 = 0$, then the vectors are linearly independent.

---

Applying the linear transformation $T$ on both sides, we obtain

$$T\left(a_1\mathrm{x} + a_2(T\mathrm{x}) + a_3(T^2\mathrm{x})\right) = T(0) = 0.$$

Using the linearity of $T$, we have

$$a_1(T\mathrm{x}) + a_2(T^2\mathrm{x}) + a_3(T^3\mathrm{x}) = 0$$

and since $T^3 = 0$, we have

$$a_1(T\mathrm{x}) + a_2(T^2\mathrm{x}) = 0 \tag{**}$$

r

---

We apply the linear transformation $T$ again, and by using the linearity as above we obtain

$$a_1(T^2\mathrm{x}) + a_2(T^3\mathrm{x}) = T(0) = 0$$

and since $T^3 = 0$, we have

$$a_1(T^2\mathrm{x}) = 0.$$

Since $T^2\mathrm{x} \neq 0$, we must have $a_1 = 0$.

---

Then from (**), we have $a_2(T^2\mathrm{x}) = 0$, and hence $a_2 = 0$ since $T^2\mathrm{x} \neq 0$.
Substituting $a_1 = a_2 = 0$ in (*), we have $a_3(T^2\mathrm{x}) = 0$, and thus $a_3 = 0$ since $T^2\mathrm{x} \neq 0$.

Therefore the solution to (*) is $a_1 = a_2 = a_3 = 0$, and hence the vectors

$$\mathrm{x}, T\mathrm{x}, T^2\mathrm{x}$$

are linearly independent.

We conclude, hence, that these vectors form a basis for $\mathbb{R}^3$.

㉟ Let $n$ be a positive integer. Let $T : \mathbb{R}^n \to \mathbb{R}$ be a non-zero linear transformation. Prove the followings.
**(a)** The nullity of $T$ is $n-1$. That is, the dimension of the nullspace of $T$ is $n-1$.
**(b)** Let $B = \{\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}\}$ be a basis of the nullspace $\mathcal{N}(T)$ of $T$. Let $\mathbf{w}$ be the $n$-dimensional vector that is not in $\mathcal{N}(T)$. Then $B' = \{\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}, \mathbf{w}\}$ is a basis of $\mathbb{R}^n$.
**(c)** Each vector $\mathbf{u} \in \mathbb{R}^n$ can be expressed as

$$\mathbf{u} = \mathbf{v} + \frac{T(\mathbf{u})}{T(\mathbf{w})}\mathbf{w}$$

for some vector $\mathbf{v} \in \mathcal{N}(T)$.

Solution.

(a) The nullity of $T$ is $n-1$.

Let $A$ be the matrix representation of the linear transformation $T:\mathbb{R}^n \to \mathbb{R}$.
Then $A$ is a $1\times n$ matrix and nonzero since $T$ is a non-zero linear transformation.

Thus, the rank of $A$ is $1$. So the rank of $T$ is $1$.
Then the rank-nullity theorem

$$\text{rank of } T + \text{ nullity of } T = n$$

yields that the nullity of $T$ is $n-1$.

(b) $B'$ is a basis of $\mathbb{R}^n$.

We claim that the $n$ vectors $\mathrm{v}_1, \cdots, \mathrm{v}_{n-1}, \mathrm{w}$ are linearly independent. Suppose that we have

$$c_1\mathrm{v}_1 + \cdots + c_{n-1}\mathrm{v}_{n-1} + c_n\mathrm{w} = 0,$$

for some $c_1, \ldots, c_n \in \mathbb{R}$.
Observe that if $c_n \neq 0$, then we have

$$\mathrm{w} = \frac{-c_1}{c_n}\mathrm{v}_1 + \cdots + \frac{-c_{n-1}}{c_n}\mathrm{v}_{n-1},$$

and thus $\mathrm{w} \in \mathrm{Span}(B) = \mathcal{N}(T)$ , a contradiction.
Hence $c_n = 0$.

---

Then the linear combination becomes

$$c_1\mathrm{v}_1 + \cdots + c_{n-1}\mathrm{v}_{n-1} = 0,$$

and since $B = \{\mathrm{v}_1, \cdots, \mathrm{v}_{n-1}\}$ is linearly independent (as being a basis), we must have
$c_1 = \cdots = c_{n-1} = 0$.
As a result, all $c_i$ are zero. So the vectors $\mathrm{v}_1, \cdots, \mathrm{v}_{n-1}, \mathrm{w}$ are linearly independent.

Since $\mathbb{R}^n$ is an $n$-dimensional vector space and $B'$ consists of $n$ linearly independent vectors, the set $B'$ is a basis of $\mathbb{R}^n$.

(c) Each $\mathrm{u} \in \mathbb{R}^n$ can be expressed as $\mathrm{u} = \mathrm{v} + \frac{T(\mathrm{u})}{T(\mathrm{w})}\mathrm{w}$ .

Let $\mathrm{u} \in \mathbb{R}^n$ . Since $B'$ isa basis of $\mathbb{R}^n$, we can write

$$\mathrm{u} = c_1\mathrm{v}_1 + \cdots + c_{n-1}\mathrm{v}_{n-1} + c_n\mathrm{w},$$

for some $c_1, \ldots, c_n \in \mathbb{R}$.
Let

$$\mathrm{u} = c_1\mathrm{v}_1 + \cdots + c_{n-1}\mathrm{v}_{n-1} + c_n\mathrm{w}.$$

Then $\mathrm{v}$ is in $\mathcal{N}(T)$ and we have

$$\mathrm{u} = \mathrm{v} + c_n \mathrm{w}. \tag{*}$$

---

We determine $c_n$ as follows.

Applying the linear transformation $T$, we obtain

$$\begin{aligned} T(\mathrm{u}) &= T(\mathrm{v} + c_n \mathrm{w}) \\ &= T(\mathrm{v}) + c_n T(\mathrm{w}) && \text{by linearity of } T \\ &= 0 + c_n T(\mathrm{w}) && \text{since } \mathrm{v} \in \mathcal{N}(T) \\ &= c_n T(\mathrm{w}). \end{aligned}$$

---

Note that since $\mathrm{w} \notin \mathcal{N}(T)$, the real number $T(\mathrm{w})$ is not zero.

Hence, we obtain

$$c_n = \frac{T(\mathrm{u})}{T(\mathrm{w})}.$$

Combining this with (*), it follows that

$$\mathrm{u} = \mathrm{v} + \frac{T(\mathrm{u})}{T(\mathrm{w})} \mathrm{w}$$

with $\mathrm{v} \in \mathcal{N}(T)$.

\end{enumerate}

\end{proof}

36 Let $V$ be the subspace of $\mathbb{R}^4$ defined by the equation

$$x_1 - x_2 + 2x_3 + 6x_4 = 0.$$

Find a linear transformation $T$ from $\mathbb{R}^3$ to $\mathbb{R}^4$ such that the null space $\mathcal{N}(T) = \{\mathbf{0}\}$ and the range $\mathcal{R}(T) = V$. Describe $T$ by its matrix $A$.

## Solution.

Any vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in V$$

can be written as

$$\mathbf{x} = \begin{bmatrix} x_2 - 2x_3 - 6x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$= x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ be the vectors appearing in the above linear combination of $\mathbf{x}$.
Then it is straightforward to see that the set $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $V$.

We define the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^4$ by

$$T(\mathbf{x}) = A\mathbf{x},$$

where

$$A = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = \begin{bmatrix} 1 & -2 & -6 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since $B$ is a basis of $V$, in particular it is a linearly independent set. Thus, the columns of $A$ is linearly independent.
It follows that the null space $\mathcal{N}(T) = \mathcal{N}(A) = \{\mathbf{0}\}$.

Also, the range of $T$ is the same as the range of $A$, which is spanned by the columns of $A$.
Thus, the range $\mathcal{R}(T) = \text{Span}(B) = V$.

By our definition of $T$, the matrix representation of $T$ is $A$.

A more explicit formula for $T$ is given by

$$T\left( \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 1 & -2 & -6 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - 2x_2 - 6x_3 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

---

**37** Let $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation

$$T(\mathbf{x}) = \text{proj}_{\mathbf{u}}\mathbf{x} = \left( \frac{\mathbf{u} \cdot \mathbf{x}}{\mathbf{u} \cdot \mathbf{u}} \right) \mathbf{u}.$$

**(a)** Calculate the null space $\mathcal{N}(T)$, a basis for $\mathcal{N}(T)$ and nullity of $T$.
**(b)** Only by using part (a) and no other calculations, find $\det(A)$, where $A$ is the matrix representation of $T$ with respect to the standard basis of $\mathbb{R}^3$.
**(c)** Calculate the range $\mathcal{R}(T)$, a basis for $\mathcal{R}(T)$ and the rank of $T$.
**(d)** Calculate the matrix $A$ representing $T$ with respect to the standard basis for $\mathbb{R}^3$.

**(e)** Let $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ be a basis for $\mathbb{R}^3$. Calculate the coordinates of $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ with respect to $B$.

(*The Ohio State University*)

---

## Proof.

### (a) Calculate the null space $\mathcal{N}(T)$, a basis for $\mathcal{N}(T)$ and nullity of $T$.

Let $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ and suppose that we have $\mathbf{x} \in \mathcal{N}(T)$.

This means that $T(\mathbf{x}) = 0$.

Thus we have

$$\left( \frac{\mathbf{u} \cdot \mathbf{x}}{\mathbf{u} \cdot \mathbf{u}} \right) \mathbf{u} = 0.$$

Since $\mathbf{u} \neq 0$, we must have

$$0 = \mathbf{u} \cdot \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x + y.$$

This yields that the null space $\mathcal{N}(T)$ consists of the vectors of the form

$$\mathbf{x} = \begin{bmatrix} x \\ -x \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad (*)$$

for any real numbers $x, z$.

Thus, we have

$$\mathcal{N}(T) = \operatorname{Span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

It follows from (*) that the vectors $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent, hence they form a basis of $\mathcal{N}(T)$.

The nullity of $T$ is the dimension of $\mathcal{N}(T)$ by definition, and hence the nullity of $T$ is $2$.

### (b) Find the determinant $\det(A)$

Note that the nullity of $A$ is the same as the nullity of $T$. So by part (a), the nullity of $A$ is $2$. This implies that the equation

$$A\mathbf{x} = 0$$

has a nonzero solution $\mathrm{x}$.

Hence the matrix $A$ is singular, and thus the determinant $\det(A) = 0$.

(c) Calculate the range $\mathcal{R}(T)$, a basis for $\mathcal{R}(T)$ and the rank of $T$.

Note that by the formula of the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$, every vector $\mathrm{x} \in \mathbb{R}^3$ is mapped to a scalar multiple of the vector $\mathrm{u}$.

Also, since $\mathrm{u}$ is mapped to $\mathrm{u}$ itself, the linear transformation $T$ is nonzero.

Hence the range is

$$\mathcal{R}(T) = \mathrm{Span}(\mathrm{u}),$$

and a basis for $\mathcal{R}(T)$ is the set $\{\mathrm{u}\}$.

The rank of $T$ is the dimension of the range $\mathcal{R}(T)$ by definition.

Thus, the rank of $T$ is $1$.

(d) Calculate the matrix $A$ representing $T$ with respect to the standard basis for $\mathbb{R}^3$.

Let

$$\mathrm{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathrm{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathrm{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

be the standard basis vectors of $\mathbb{R}^3$.

We calculate

$$T(\mathrm{e}_1) = \left( \frac{\mathrm{u} \cdot \mathrm{e}_1}{\mathrm{u} \cdot \mathrm{u}} \right) \mathrm{u} = \frac{1}{2} \mathrm{u} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$T(\mathrm{e}_2) = \left( \frac{\mathrm{u} \cdot \mathrm{e}_2}{\mathrm{u} \cdot \mathrm{u}} \right) \mathrm{u} = \frac{1}{2} \mathrm{u} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$T(\mathrm{e}_3) = \left( \frac{\mathrm{u} \cdot \mathrm{e}_3}{\mathrm{u} \cdot \mathrm{u}} \right) \mathrm{u} = \frac{0}{2} \mathrm{u} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus, the matrix representation $A$ of $T$ is

$$\begin{aligned} A &= [T(\mathrm{e}_1) \quad T(\mathrm{e}_2) \quad T(\mathrm{e}_3)] \\ &= \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

### (e) Calculate the coordinate vector

Let $\mathrm{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and let

$$[\mathrm{v}]_B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

be the coordinate vector of $\mathrm{v}$ with respect to the basis $B$.

This means that we have

$$\mathrm{v} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

To determine $a, b, c$, we solve the system

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathrm{v}.$$

The augmented matrix of the system is

$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & x \\ 0 & 1 & -1 & y \\ 0 & 0 & 1 & z \end{array}\right].$$

Applying the elementary row operations, we obtain

$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & x \\ 0 & 1 & -1 & y \\ 0 & 0 & 1 & z \end{array}\right] \xrightarrow{R_1+R_2} \left[\begin{array}{rrr|r} 1 & 0 & -1 & x+y \\ 0 & 1 & -1 & y \\ 0 & 0 & 1 & z \end{array}\right]$$

$$\xrightarrow[R_2+R_3]{R_1+R_3} \left[\begin{array}{rrr|r} 1 & 0 & 0 & x+y+x \\ 0 & 1 & 0 & y+z \\ 0 & 0 & 1 & z \end{array}\right].$$

Hence we have

$$a = x+y+z, b = y+z, c = z,$$

and the coordinate vector of $\mathrm{v}$ with respect to the basis $B$ is

$$[\mathrm{v}]_B = \begin{bmatrix} x+y+z \\ y+z \\ z \end{bmatrix}.$$

㊳ We fix a nonzero vector $\mathbf{a}$ in $\mathbb{R}^3$ and define a map $T : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$T(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$$

for all $\mathbf{v} \in \mathbb{R}^3$. Here the right-hand side is the cross product of $\mathbf{a}$ and $\mathbf{v}$.
**(a)** Prove that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation.
**(b)** Determine the eigenvalues and eigenvectors of $T$.

### Proof.

#### (a) Prove that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation.

To prove that $T$ is a linear transformation, we need to verify the following equalities for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ and $c \in \mathbb{R}$:

㊴ $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

㊵ $T(c\mathbf{v}) = cT(\mathbf{v})$.

These follow from the basic properties of cross products as follows.

We have

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= \mathbf{a} \times (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{a} \times \mathbf{u} + \mathbf{a} \times \mathbf{v} && \text{the cross product is distributive} \\ &= T(\mathbf{u}) + T(\mathbf{v}). \end{aligned}$$

As the cross product is compatible with scalar multiplication, we also have

$$T(c\mathbf{v}) = \mathbf{a} \times (c\mathbf{v}) = c(\mathbf{a} \times \mathbf{v}) = cT(\mathbf{v}).$$

Therefore $T$ is a linear transformation.

#### (b) Determine the eigenvalues and eigenvectors of $T$.

Let $\lambda$ be an eigenvalue of the linear transformation $T$ and let $\mathbf{x}$ be an eigenvector corresponding to $\lambda$. Then we have

$$T(\mathbf{x}) = \mathbf{a} \times \mathbf{x} = \lambda \mathbf{x}.$$

We take the dot product and obtain

$$(a \times \mathbf{x}) \cdot \mathbf{x} = \lambda \mathbf{x} \cdot \mathbf{x} = \lambda \|\mathbf{x}\|^2.$$

Note that the vector $\mathbf{a} \times \mathbf{x}$ and $\mathbf{x}$ is orthogonal, hence the left-hand side of the above equality is zero.
On the other hand, since $\mathbf{x}$ is an eigenvector, it is a nonzero vector. Hence the length $\|\mathbf{x}\| \neq 0$.
It follows that we have $0 = \lambda \|\mathbf{x}\|^2$ with $\|\mathbf{x}\| \neq 0$, and thus $\lambda = 0$.
Hence the eigenvalue of $T$ is $\lambda = 0$.

This yields that we have $\mathbf{a} \times \mathbf{x} = 0 \cdot \mathbf{x} = \mathbf{0}$ for any eigenvector $\mathbf{x}$.

Note that the cross product is zero indicates that the vectors $\mathbf{a}$ and $\mathbf{x}$ are parallel.

Hence $\mathbf{x} = r\mathbf{a}$ for a nonzero $r \in \mathbb{R}$.

In summary, the eigenvalue of $T$ is $0$ and corresponding eigenvectors are $r\mathbf{a}$ for any nonzero real number $r \in \mathbb{R}$.

## Comment.

Let $\{\mathbf{e}_1, \mathbf{e}_2 \mathbf{e}_3\}$ be the standard basis for $\mathbb{R}^3$ and write $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$.

Then we have

$$T(\mathbf{e}_1) = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix}$$
$$T(\mathbf{e}_2) = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix}$$
$$T(\mathbf{e}_3) = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix}.$$

It follows that the matrix representation of the linear transformation $T$ with respect to the standard basis is given by

$$A = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}.$$

Note that the matrix $A$ is a real skew-symmetric matrix.

We know that each eigenvalue of a skew-symmetric matrix is either $0$ or purely imaginary.

(See the post "".)

Also if the degree of the matrix is odd, it has $0$ as an eigenvalue.

---

41 Determine all linear transformations of the $2$-dimensional $x$-$y$ plane $\mathbb{R}^2$ that take the line $y = x$ to the line $y = -x$.

---

## Solution.

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps the line $y = x$ to the line $y = -x$.

Note that the linear transformation $T$ is completely determined if the values of $T$ on basis vectors of the vector space $\mathbb{R}^2$ are known.

Let

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

be a basis of $\mathbb{R}^2$.

---

The reason of this choice is as follows.

Since we know that $T$ takes the line $y = x$ to the line $y = -x$, the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is mapped into some point on the line $y = -x$.

This is how we chose the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The other basis vector could be any vector that is not a multiple of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and we just chose the simple vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

---

Let

$$T\left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} a \\ b \end{bmatrix}$$

for some $a, b \in \mathbb{R}$.

Since we know that the vector $T\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$ is on the line $y = -x$, let

$$T\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} c \\ -c \end{bmatrix}$$

for some $c \in \mathbb{R}$.

---

We now find a formula for this linear transformation $T$.

Let $\begin{bmatrix} x \\ y \end{bmatrix}$ be an arbitrary vector in the plane $\mathbb{R}^2$.

We express this vector as a linear combination of basis vectors:

$$\begin{bmatrix} x \\ y \end{bmatrix} = (x - y) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Then we have

$$\begin{aligned}
&T\left( \begin{bmatrix} x \\ y \end{bmatrix} \right) \\
&= T\left( (x - y) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \\
&= (x - y) T\left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) + y T\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) && \text{since } T \text{ is a linear transformation} \\
&= (x - y) \begin{bmatrix} a \\ b \end{bmatrix} + y \begin{bmatrix} c \\ -c \end{bmatrix} \\
&= \begin{bmatrix} ax + (c - a)y \\ bx - (c + b)y \end{bmatrix}.
\end{aligned}$$

We conclude that any linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that takes the line $y = x$ to the line $y = -x$ is of the form

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} ax + (c-a)y \\ bx - (c+b)y \end{bmatrix}$$

for some $a, b, c \in \mathbb{R}$.

### Remark.

Remark that if $c = 0$, then all the points on the line $y = x$ are mapped into the origin, which is on the line $y = -x$.
If we want to avoid this degenerate case, we need to assume that $c \neq 0$.

(42) Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be the function that maps each vector in $\mathbb{R}^2$ to its reflection with respect to $x$-axis. Determine the formula for the function $F$ and prove that $F$ is a linear transformation.

### Solution 1.

Let $\begin{bmatrix} x \\ y \end{bmatrix}$ be an arbitrary vector in $\mathbb{R}^2$.
Its reflection with respect to $x$-axis is the vector $\begin{bmatrix} x \\ -y \end{bmatrix}$.
Thus the formula for the function $F$ is given by

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ -y \end{bmatrix}.$$

---

Next, we prove that the function $F$ is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$.
We need to verify the following two properties: for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ and $c \in \mathbb{R}$, we have

(43) $F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v})$

(44) $F(c\mathbf{u}) = cF(\mathbf{u})$.

Let $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} x' \\ y' \end{bmatrix}$.
Then we have

$$\begin{aligned}
F(\mathbf{u}+\mathbf{v}) &= F\left(\begin{bmatrix} x+x' \\ y+y' \end{bmatrix}\right) \\
&= \begin{bmatrix} x+x' \\ -(y+y') \end{bmatrix} && \text{by the formula for } F \\
&= \begin{bmatrix} x \\ -y \end{bmatrix} + \begin{bmatrix} x' \\ -y' \end{bmatrix} \\
&= F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) + F\left(\begin{bmatrix} x' \\ y' \end{bmatrix}\right) && \text{by the formula for } F \\
&= F(\mathbf{u}) + F(\mathbf{v}).
\end{aligned}$$

This prove property (a).

---

Next, we have

$$\begin{aligned}
F(c\mathbf{u}) &= F\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right) \\
&= \begin{bmatrix} cx \\ -(cy) \end{bmatrix} && \text{by the formula for } F \\
&= c\begin{bmatrix} x \\ -y \end{bmatrix} \\
&= cF(\mathbf{u}) && \text{by the formula for } F.
\end{aligned}$$

Hence property (b) is verified.

Therefore, the function $F$ is a linear transformation.

## Solution 2.

We give another proof that $F$ is a linear transformation.

From the formula that we obtained in Solution 1, we see that

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

It follows that the function $F$ is the matrix multiplication by $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

As the matrix multiplication function is always a linear transformation, we conclude that $F$ is a linear transformation.

---

Observe that the matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is the matrix representation of the linear transformation $F$.

That is,

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = [T(\mathbf{e}_1), T(\mathbf{e}_2],$$

where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are standard basis vectors of $\mathbb{R}^2$.

---

(45) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation of the $2$-dimensional vector space $\mathbb{R}^2$ (the $x$-$y$-plane) to itself which is the reflection across a line $y = mx$ for some $m \in \mathbb{R}$. Then find the matrix representation of the linear transformation $T$ with respect to the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ of $\mathbb{R}^2$, where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

### Solution.

Let $A$ be the matrix representation of $T$ with respect to the standard basis $B$.

Observe that each vector on the line $y = mx$ does not move under the linear transformation $T$.

Since the vector $\begin{bmatrix} 1 \\ m \end{bmatrix}$ is on the line $y = mx$, it follows that

$$A \begin{bmatrix} 1 \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ m \end{bmatrix}. \qquad (*)$$

Note that if $m \neq 0$ then the line $y = \frac{-1}{m} x$ is perpendicular to the line $y = mx$ at the origin.

If $m = 0$, then the line $x = 0$ is perpendicular to the line $y = 0$ at the origin.

In either case the vector $\begin{bmatrix} -m \\ 1 \end{bmatrix}$ is on the perpendicular line.

Thus, by the reflection across the line $y = mx$, this vector is mapped to $\begin{bmatrix} m \\ -1 \end{bmatrix}$.

That is, we have

$$A \begin{bmatrix} -m \\ 1 \end{bmatrix} = \begin{bmatrix} m \\ -1 \end{bmatrix}. \qquad (**)$$

It follows from (*) and (**) that

$$A \begin{bmatrix} 1 & -m \\ m & 1 \end{bmatrix} = \left[ A \begin{bmatrix} 1 \\ m \end{bmatrix} \quad A \begin{bmatrix} -m \\ 1 \end{bmatrix} \right] = \begin{bmatrix} 1 & m \\ m & -1 \end{bmatrix}.$$

The determinant of the matrix $\begin{bmatrix} 1 & -m \\ m & 1 \end{bmatrix}$ is $1 + m^2 \neq 0$, hence it is invertible.

(Note that since column vectors are nonzero orthogonal vectors, we knew it is invertible.)

The inverse matrix is

$$\begin{bmatrix} 1 & -m \\ m & 1 \end{bmatrix}^{-1} = \frac{1}{1 + m^2} \begin{bmatrix} 1 & m \\ -m & 1 \end{bmatrix}.$$

Therefore, we have

$$
\begin{aligned}
A &= \begin{bmatrix} 1 & m \\ m & -1 \end{bmatrix} \begin{bmatrix} 1 & -m \\ m & 1 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 1 & m \\ m & -1 \end{bmatrix} \cdot \frac{1}{1+m^2} \begin{bmatrix} 1 & m \\ -m & 1 \end{bmatrix} \\
&= \frac{1}{1+m^2} \begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.
\end{aligned}
$$

In summary, the matrix representation $A$ of the linear transformation $T$ across the line $y = mx$ with respect to the standard basis is

$$A = \frac{1}{1+m^2} \begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.$$

## Comment.

The matrix representation $A$ of a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ with respect to the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ of $\mathbb{R}^2$ is given by

$$A = [T(\mathbf{e}_1) \quad T(\mathbf{e}_2)].$$

So we can technically find the matrix $A$ by finding $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$, which can be found by elementary plane geometry.

However, it is easier to determine the outputs of $T$ for vectors on the line $y = mx$ and vectors on the perpendicular line.
That why we didn't compute the vectors $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ directly in the above solution.

---

46 Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation and suppose that its matrix representation with respect to the standard basis is given by the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 4 & 0 & 5 \end{bmatrix}$.

**(a)** Prove that the linear transformation $T$ sends points on the $x$-$z$ plane to points on the $x$-$z$ plane.
**(b)** Prove that the restriction of $T$ on the $x$-$z$ plane is a linear transformation.
**(c)** Find the matrix representation of the linear transformation obtained in part (b) with respect to the standard basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ of the $x$-$z$ plane.

## Proof.

(a) Prove that the linear transformation $T$ sends points on the $x$-$z$ plane to points on the $x$-$z$ plane.

Each point on the $x$-$z$ plane is of the form

$$\begin{bmatrix} x \\ 0 \\ z \end{bmatrix}$$

for some $x, z \in \mathbb{R}$.

We have

$$\begin{aligned} T\left(\begin{bmatrix} x \\ 0 \\ z \end{bmatrix}\right) &= A\begin{bmatrix} x \\ 0 \\ z \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 4 & 0 & 5 \end{bmatrix}\begin{bmatrix} x \\ 0 \\ z \end{bmatrix} \\ &= \begin{bmatrix} x+2z \\ 0 \\ 4x+5z \end{bmatrix}. \end{aligned}$$

Since the $y$-coordinate of the last vector is $0$, and thus the output vector lies in the $x$-$z$ plane.

> (b) Prove that the restriction of $T$ on the $x$-$z$ plane is a linear transformation.

Let $V$ be the $x$-$z$ plane in $\mathbb{R}^3$.
Then $V$ is a subspace of the vector space $\mathbb{R}^3$.

In part (a), we showed that the restriction of $T$ on $V$ is given by the formula

$$T\left(\begin{bmatrix} x \\ 0 \\ z \end{bmatrix}\right) = \begin{bmatrix} x+2z \\ 0 \\ 4x+5z \end{bmatrix}. \tag{*}$$

We abuse the notation and write this restriction as $T : V \to V$.
(The precise notation is $T|_V : V \to V$.)
Let

$$\begin{bmatrix} x_1 \\ 0 \\ z_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ 0 \\ z_2 \end{bmatrix}$$

be arbitrary vectors in $V$ and let $r \in \mathbb{R}$ be an arbitrary real number.
Then we have

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ 0 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ 0 \\ z_2 \end{bmatrix}\right) &= T\left(\begin{bmatrix} x_1+x_2 \\ 0 \\ z_1+z_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} (x_1+x_2)+2(z_1+z_2) \\ 0 \\ 4(x_1+x_2)+5(z_1+z_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1+2z_1 \\ 0 \\ 4x_1+5z_1 \end{bmatrix} + \begin{bmatrix} x_2+2z_2 \\ 0 \\ 4x_2+5z_2 \end{bmatrix} \\ &= T\left(\begin{bmatrix} x_1 \\ 0 \\ z_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ 0 \\ z_2 \end{bmatrix}\right) \end{aligned}$$

and

$$\begin{aligned} T\left(r\begin{bmatrix} x_1 \\ 0 \\ z_1 \end{bmatrix}\right) &= T\left(\begin{bmatrix} rx_1 \\ 0 \\ rz_1 \end{bmatrix}\right) \\ &= \begin{bmatrix} (rx_1)+2(rz_1) \\ 0 \\ 4(rx_1)+5(z_1) \end{bmatrix} \\ &= r\begin{bmatrix} x_1+2z_1 \\ 0 \\ 4x_1+5z_1 \end{bmatrix} \\ &= rT\left(\begin{bmatrix} x_1 \\ 0 \\ z_1 \end{bmatrix}\right). \end{aligned}$$

It follows that the restriction $T : V \to V$ is a linear transformation.

> (c) Find the matrix representation of the linear transformation obtained in part (b) with respect to the standard basis

Let $B = \{\mathbf{u}, \mathbf{v}\}$ be the standard basis of the $x$-$z$ plane:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

It follows from formula (*) that we have

$$T(\mathbf{u}) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \mathbf{u} + 4\mathbf{v}$$

and

$$T(\mathbf{v}) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix} = 2\mathbf{u} + 5\mathbf{v}.$$

Thus the coordinate vectors with respect to the basis $B$ are

$$[T(\mathbf{u})]_B = \begin{bmatrix} 1 \\ 4 \end{bmatrix}_B, [T(\mathbf{v})]_B = \begin{bmatrix} 2 \\ 5 \end{bmatrix}_B,$$

and the matrix representation of the linear transformation $T : V \to V$ with respect to the standard basis $B = \{\mathbf{u}, \mathbf{v}\}$ is

$$[\,[T(\mathbf{u})]_B, [T(\mathbf{v})]_B\,] = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}.$$

㊼ Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by the formula

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_2 - 2x_3 \\ 2x_1 + 3x_2 \\ x_2 + x_3 \end{bmatrix}.$$

Determine whether $T$ is an isomorphism and if so find the formula for the inverse linear transformation $T^{-1}$.

## Solution.

Let $B = \{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ be the standard basis of $\mathbb{R}^3$, where

$$\mathrm{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathrm{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathrm{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We determine the matrix representation $[T]_B$ of $T$ with respect to the basis $B$.

Since we have

$$T(\mathrm{e}_1) = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, T(\mathrm{e}_2) = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}, T(\mathrm{e}_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix},$$

we have

$$[T]_B = [\, T(\mathrm{e}_1) \quad T(\mathrm{e}_2) \quad T(\mathrm{e}_3)] = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

This matrix is invertible and the inverse matrix is given by

$$[T]_B^{-1} = \begin{bmatrix} 3 & -1 & -6 \\ -2 & 1 & 4 \\ -2 & 1 & 3 \end{bmatrix}.$$

(See the post Find the Inverse Matrices if Matrices are Invertible by Elementary Row Operations for details of how to find the inverse matrix of this matrix.)

This implies that the matrix $T$ is an isomorphism.

---

Observe that we have $[T]_B^{-1} = [T^{-1}]_B$.

Thus, we obtain

$$T^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = [T^{-1}]_B \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & -1 & -6 \\ -2 & 1 & 4 \\ -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 3x_1 - x_2 - 6x_3 \\ -2x_1 + x_2 + 4x_3 \\ -2x_1 + x_2 + 3x_3 \end{bmatrix}.$$

---

In summary, the formula for the inverse linear transformation $T^{-1}$ is given by

$$T^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 - x_2 - 6x_3 \\ -2x_1 + x_2 + 4x_3 \\ -2x_1 + x_2 + 3x_3 \end{bmatrix}.$$

48 Let $\mathbb{R}^n$ be an inner product space with inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is called **orthogonal transformation** if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, it satisfies

$$\langle T(\mathbf{x}), T(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Prove that if $T : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, then $T$ is an isomorphism.

## Proof 1.

As $T$ is a linear transformation from $\mathbb{R}^n$ to itself, it suffices to show that $T$ is an injective linear transformation.

Suppose that $T(\mathrm{x}) = T(\mathrm{y})$ for $\mathrm{x}, \mathrm{y} \in \mathbb{R}^n$ .
We show that $\mathrm{x} = \mathrm{y}$ .

We have

$$\begin{aligned} &\|\mathrm{x} - \mathrm{y}\|^2 \\ &= (\mathrm{x} - \mathrm{y})^{\mathrm{T}}(\mathrm{x} - \mathrm{y}) \\ &= (\mathrm{x}^{\mathrm{T}} - \mathrm{y}^{\mathrm{T}})(\mathrm{x} - \mathrm{y}) \\ &= \mathrm{x}^{\mathrm{T}}\mathrm{x} - \mathrm{x}^{\mathrm{T}}\mathrm{y} - \mathrm{y}^{\mathrm{T}}\mathrm{x} + \mathrm{y}^{\mathrm{T}}\mathrm{y} \\ &= \langle \mathrm{x}, \mathrm{x} \rangle - \langle \mathrm{x}, \mathrm{y} \rangle - \langle \mathrm{y}, \mathrm{x} \rangle + \langle \mathrm{y}, \mathrm{y} \rangle \\ &= \langle T(\mathrm{x}), T(\mathrm{x}) \rangle - \langle T(\mathrm{x}), T(\mathrm{y}) \rangle - \langle T(\mathrm{y}), T(\mathrm{x}) \rangle + \langle T(\mathrm{y}), T(\mathrm{y}) \rangle \\ &(\text{since } T \text{ is an orthogonal transformation}) \\ &= \langle T(\mathrm{x}), T(\mathrm{x}) \rangle - \langle T(\mathrm{x}), T(\mathrm{x}) \rangle - \langle T(\mathrm{x}), T(\mathrm{x}) \rangle + \langle T(\mathrm{x}), T(\mathrm{x}) \rangle \\ &(\text{since } T(\mathrm{x}) = T(\mathrm{y})) \\ &= 0. \end{aligned}$$

It follows that $\|\mathrm{x} - \mathrm{y}\| = 0$ and hence $\mathrm{x} = \mathrm{y}$ .

This proves that $T : \mathbb{R}^n \to \mathbb{R}^n$ is injective.

As $T$ is an injective linear transformation from the $n$-dimensional vector space $\mathbb{R}^n$ to itself, it is also surjective, and thus $T$ is an isomorphism.

## Proof 2.

Recall that the linear transformation $T$ is injective if and only if the null space $\mathcal{N}(T) = \{0\}$, that is, $T(\mathrm{x}) = 0$ implies that $\mathrm{x} = 0$.

We use this fact to show that $T$ is injective.
Suppose that $T(\mathrm{x}) = 0$.
Then we have

$$\begin{aligned} \|\mathrm{x}\|^2 &= \langle \mathrm{x}, \mathrm{x} \rangle \\ &= \langle T(\mathrm{x}), T(\mathrm{x}) \rangle && \text{as } T \text{ is orthogonal} \\ &= \langle 0, 0 \rangle = 0 && \text{as } T(\mathrm{x}) = 0. \end{aligned}$$

It follows that the length $\|\mathrm{x}\| = 0$, and hence $\mathrm{x} = 0$.
This proves that the null space $\mathcal{N}(T) = \{0\}$ and $T$ is injective.

As $T$ is an injective linear transformation from $\mathbb{R}^n$ to itself, it is also surjective, and hence $T$ is an isomorphism.

# Linear Transformation Between Vector Spaces

**Definition**

Let $V$ and $W$ be vector spaces over a scalar field $K$.

① A function $T : V \to W$ is called a **linear transformation** if $T$ satisfies the following two linearity conditions: For any $\mathbf{x}, \mathbf{y} \in V$ and $c \in K$, we have

1. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$
2. $T(c\mathbf{x}) = cT(\mathbf{x})$

② The **nullspace** $\mathcal{N}(T)$ of a linear transformation $T : V \to W$ is

$$\mathcal{N}(T) = \{\mathbf{x} \in V \mid T(\mathbf{x}) = \mathbf{0}\}.$$

③ The **nullity** of $T$ is the dimension of $\mathcal{N}(T)$.

④ The **range** $\mathcal{R}(T)$ of a linear transformation $T : V \to W$ is

$$\mathcal{R}(T) = \{\mathbf{y} \in W \mid \mathbf{y} = T(\mathbf{x}) \text{ for some } \mathbf{x} \in V\}.$$

⑤ The **rank** of $T$ is the dimension of $\mathcal{R}(T)$.

**Summary**

Let $T : V \to W$ be a linear transformation.

① $T$ maps the zero vector $\mathbf{0}_V$ of $V$ to the zero vector $\mathbf{0}_W$ of $W$. That is, $T(\mathbf{0}_V) = \mathbf{0}_W$.

## Problems

① For an integer $n > 0$, let $\mathrm{P}_n$ be the vector space of polynomials of degree at most $n$. The set $B = \{1, x, x^2, \cdots, x^n\}$ is a basis of $\mathrm{P}_n$, called the standard basis. Let $T : \mathrm{P}_n \to \mathrm{P}_{n+1}$ be the map defined by, for $f \in \mathrm{P}_n$, $T(f)(x) = xf(x)$. Prove that $T$ is a linear transformation, and find its range and nullspace.

### Solution.

#### $T$ is a linear transformation

We must show that for polynomials $f, g \in \mathrm{P}_n$ and scalar $c \in \mathbb{R}$, the function $T$ satisfies $T(f+g) = T(f) + T(g)$ and $T(cf) = cT(f)$. The first is checked using associativity of multiplication:

$$T(f+g)(x) = x(f+g)(x) = xf(x) + xg(x) = T(f)(x) + T(g)(x).$$

Similarly,

$$T(cf)(x) = x(cf)(x) = cxf(x) = cT(f)(x).$$

### The nullspace of $T$

The nullspace of $T$ is the set of polynomials $f(x)$ such that $T(f) = 0$. That is, $xf(x) = 0$. But this product can only be $0$ if one of the terms being multiplied is $0$. Because $x \neq 0$, we must have that $f(x) = 0$. Thus the nullspace is the trivial vector subspace $\{0\}$.

### The range of $T$

The range of $T$ is the set of polynomials of the form $xf(x)$ for an arbitrary polynomial $f \in \mathrm{P}_n$. We can calculate this by calculating $T$ for a basis of $\mathrm{P}_n$. Let $B = \{1, x, x^2, \cdots, x^n\}$ be a basis of $\mathrm{P}_n$.

We see that $T(x^i) = x^{i+1}$ for $0 \leq i \leq n$, and so the image of the basis is the set

$$T(B) = \{x, x^2, x^3, \cdots, x^{n+1}\}.$$

Then the range of $T$ must be the span of this set. Specifically,

$$\mathcal{R}(T) = \mathrm{Span}(T(B)) = \mathrm{Span}(x, x^2, x^3, \cdots, x^{n+1}).$$

---

2 Let $\mathrm{P}_n$ be the vector space of polynomials of degree at most $n$. The set $B = \{1, x, x^2, \cdots, x^n\}$ is a basis of $\mathrm{P}_n$, called the standard basis. Let $T : \mathrm{P}_4 \to \mathrm{P}_4$ be the map defined by, for $f \in \mathrm{P}_4$, $T(f)(x) = f(x) - x - 1$. Determine if $T(x)$ is a linear transformation. If it is, find the matrix representation of $T$ relative to the standard basis of $\mathrm{P}_4$.

### Solution 1.

We claim that $T$ is not a linear transformation.

To see this, consider $f(x) = 1 + x$. Then

$$T(f)(x) = 1 + x - x - 1 = 0.$$

On the other hand,

$$T(1) + T(x) = (1 - x - 1) + (x - x - 1) = -x - 1.$$

Because $T(1 + x) \neq T(1) + T(x)$, we see that $T$ is not a linear transformation.

### Solution 2.

Recall the fact that any linear transformation maps the zero vector to the zero vector.
The zero vector in $\mathrm{P}_n$ is the zero polynomial $\theta(x) = 0$.

Then we have

$$T(\theta)(x) = \theta(x) - x - 1 = -x - 1 \neq \theta(x).$$

Thus, the map $T$ does not send the zero vector $\theta(x)$ to the zero vector $\theta(x)$.
Hence, $T$ is not a linear transformation.

---

③ Let $\mathrm{P}_n$ be the vector space of polynomials of degree at most $n$. The set $B = \{1, x, x^2, \cdots, x^n\}$ is a basis of $\mathrm{P}_n$, called the **standard basis**. Let $T : \mathrm{P}_3 \to \mathrm{P}_5$ be the map defined by, for $f \in \mathrm{P}_3, T(f)(x) = (x^2 - 2)f(x)$. Determine if $T(x)$ is a linear transformation. If it is, find the matrix representation of $T$ relative to the standard basis of $\mathrm{P}_3$ and $\mathrm{P}_5$.

## Solution.

We must check that $T$ satisfies the two axioms for linear transformations. Suppose $f, g \in \mathrm{P}_3$. Then

$$T(f+g)(x) = (x^2-2)(f+g)(x) = (x^2-2)f(x) + (x^2-2)g(x) = T(f)(x) + T(g)(x).$$

Second, if $c \in \mathbb{R}$ then

$$T(cf)(x) = (x^2-2)(cf)(x) = c(x^2-2)f(x) = cT(f)(x).$$

The two axioms are satisfied, and so $T$ is a linear transformation.

---

Now we find its matrix representation relative to the standard basis $B = \{1, x, x^2, x^3\}$ of $\mathrm{P}_3$ and the standard basis $C = \{1, x, x^2, x^3, x^4, x^5\}$. To do this, for every $f \in B$ we calculate the coordinate vector of $T(f)$ relative to the basis $C$.

For example, we find that $T(1) = x^2 - 2$. The coordinate vector for this element, relative to $C$, is

$$[x^2-2]_C = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Similarly, we find that $T(x) = x^3 - 2x$. The coordinate vector for this element, relative to $C$,

$$[x^3-2x]_C = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Performing this process for the rest of the elements of $B$, we get

$$[T(x^2)]_C = [x^4-2x^2]_C = \begin{bmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad [T(x^3)]_C = [x^5-2x^3]_C = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}.$$

And now putting these four column vectors in order, we get the matrix representation for $T$:

$$[T]_B^C = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

④ Let $P_3$ be the vector space of polynomials of degree $3$ or less with real coefficients.
**(a)** Prove that the differentiation is a linear transformation. That is, prove that the map $T : P_3 \to P_3$ defined by

$$T\,(\,f(x)\,) = \frac{d}{dx} f(x)$$

for any $f(x) \in P_3$ is a linear transformation.
**(b)** Let $B = \{1, x, x^2, x^3\}$ be a basis of $P_3$. With respect to the basis $B$, find the matrix representation of the linear transformation $T$ in part (a).

---

### Proof.

### (a) Prove that the differentiation is a linear transformation.

Let $f(x), g(x) \in P_3$ . By the basic properties of differentiations, we have

$$\begin{aligned} T\,(\,f(x) + g(x)\,) &= \frac{d}{dx}(\,f(x) + g(x)\,) \\ &= \frac{d}{dx}(\,f(x)\,) + \frac{d}{dx}(\,g(x)\,) \\ &= T\,(\,f(x)\,) + T\,(\,g(x)\,)\,. \end{aligned}$$

For $f(x) \in P_3$ and $r \in \mathbb{R}$, we also have

$$\begin{aligned} T\,(\,rf(x)\,) &= \frac{d}{dx}(\,rf(x)\,) \\ &= r\frac{d}{dx}(\,f(x)\,) \\ &= rT\,(\,f(x)\,)\,. \end{aligned}$$

Therefore, the map $T$ is a linear transformation.

### (b) Find the matrix representation of the linear transformation $T$ .

We compute

$$T(1) = \frac{d}{dx}(1) = 0, \qquad T(x) = \frac{d}{dx}(x) = 1$$

$$T(x^2) = \frac{d}{dx}(x^2) = 2x, \qquad T(x^3) = \frac{d}{dx}(x^3) = 3x^2.$$

It follows that the coordinate vectors are

$$[T(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad [T(x)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$[T(x^2)]_B = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \qquad [T(x^3)]_B = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}.$$

Thus the matrix representation of the linear transformation $T$ is given by

$$\left[ [T(1)]_B, [T(x)]_B, [T(x^2)]_B, [T(x^3)]_B \right] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Remark that we also write this as

$$\begin{aligned} &[T(1) \quad T(x) \quad T(x^2) \quad T(x^3)] \\ &= [1 \quad x \quad x^2 \quad x^3] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

## Related Question.

Let $A$ be the matrix representation obtained in part (b).

It is easy to see that $A^4$ is the zero matrix.

We say that $A$ is a **nilpotent matrix**.

There is a theoretical explanation behind this.

Check out the post "Differentiating Linear Transformation is Nilpotent" for an explanation.

⑤ Let $P_2(\mathbb{R})$ be the vector space over $\mathbb{R}$ consisting of all polynomials with real coefficients of degree $2$ or less. Let $B = \{1, x, x^2\}$ be a basis of the vector space $P_2(\mathbb{R})$. For each linear transformation $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ defined below, find the matrix representation of $T$ with respect to the basis $B$. For $f(x) \in P_2(\mathbb{R})$, define $T$ as follows.

**(a)**

$$T(f(x)) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) - 3\frac{\mathrm{d}}{\mathrm{d}x} f(x)$$

**(b)**

$$T(f(x)) = \int_{-1}^{1} (t-x)^2 f(t)\,\mathrm{d}t$$

**(c)**

$$T(f(x)) = e^x \frac{\mathrm{d}}{\mathrm{d}x}(e^{-x} f(x))$$

### Hint.

To find the matrix, we just need to compute $T(1), T(x), T(x^2)$ and read the coefficients of $1, x, x^2$.

### Solution.

#### (a) $T\big(f(x)\big) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) - 3\frac{\mathrm{d}}{\mathrm{d}x} f(x)$

Since the derivative the constant function $1$ is zero, we have

$$T(1) = 0.$$

Since the second derivative of the function $x$ is zero, we have

$$T(x) = -3.$$

Finally, we calculate

$$T(x^2) = 2 - 6x.$$

There with respect to the basis $\{1, x, x^2\}$, the matrix for $T$ is

$$\begin{bmatrix} 0 & -3 & 2 \\ 0 & 0 & -6 \\ 0 & 0 & 0 \end{bmatrix}.$$

#### (b) $T\big(f(x)\big) = \int_{-1}^{1} (t - x)^2 f(t)\,\mathrm{d}t$

We first compute

$$\begin{aligned} T\big(f(x)\big) &= \int_{-1}^{1} (t^2 - 2tx - +x^2) f(t)\,\mathrm{d}t \\ &= \left( \int_{-1}^{1} t^2 f(t)\,\mathrm{d}t \right) - 2\left( \int_{-1}^{1} t f(t)\,\mathrm{d}t \right) x + \left( \int_{-1}^{1} f(t)\,\mathrm{d}t \right) x^2. \end{aligned}$$

Then we have

$$\begin{aligned} T(1) &= \frac{2}{3} + 2x^2 \\ T(x) &= -\frac{4}{x}x \\ T(x^2) &= \frac{2}{5} + \frac{2}{3}x^2 \end{aligned}$$

Here we calculated the integrals

$$\int_{-1}^{1} 1\,\mathrm{d}t = [t]_{-1}^{1} = 2, \quad \int_{-1}^{1} t\,\mathrm{d}t = \left[\frac{1}{2}t^2\right]_{-1}^{1} = 0,$$

$$\int_{-1}^{1} t^2\,\mathrm{d}t = \left[\frac{1}{3}t^3\right]_{-1}^{1} = \frac{2}{3}, \quad \int_{-1}^{1} t^3\,\mathrm{d}t = \left[\frac{1}{4}t^4\right]_{-1}^{1} = 0, \quad \int_{-1}^{1} t^4\,\mathrm{d}t = \left[\frac{1}{5}t^5\right]_{-1}^{1} = \frac{2}{5}.$$

Therefore, the matrix for the linear transformation with respect to the basis $B = \{1, x, x^2\}$ is

$$\begin{bmatrix} \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & -\frac{4}{3} & 0 \\ 2 & 0 & \frac{2}{3} \end{bmatrix}.$$

### (c) $T(f(x)) = e^x \frac{\mathrm{d}}{\mathrm{d}x}(e^{-x} f(x))$

Note that by the product rule for derivatives,

$$\begin{aligned} T(f(x)) &= e^x \left( -e^{-x} f(x) + e^{-x} \frac{\mathrm{d}}{\mathrm{d}x} f(x) \right) \\ &= -f(x) + \frac{\mathrm{d}}{\mathrm{d}x} f(x). \end{aligned}$$

Thus we have

$$\begin{aligned} T(1) &= -1 \\ T(x) &= -x + 1 \\ T(x^2) &= -x^2 + 2x. \end{aligned}$$

Hence the matrix for the linear transformation $T$ with respect to the basis $B$ is

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{bmatrix}.$$

⑥ Let $U$ and $V$ be vector spaces over a scalar field $\mathbb{F}$. Define the map $T : U \to V$ by $T(\mathbf{u}) = \mathbf{0}_V$ for each vector $\mathbf{u} \in U$.

**(a)** Prove that $T : U \to V$ is a linear transformation. (Hence, $T$ is called the **zero transformation**.)

**(b)** Determine the null space $\mathcal{N}(T)$ and the range $\mathcal{R}(T)$ of $T$.

## Proof.

### (a) Prove that $T : U \to V$ is a linear transformation.

Let $\mathrm{u}_1, \mathrm{u}_2 \in U$ and $r$ be a scalar, that is, $r \in \mathbb{F}$.

It follows from the definition of $T$ that

$$\begin{aligned} T(\mathrm{u}_1) &= 0_V, \quad & T(\mathrm{v}_2) &= 0_V, \\ T(\mathrm{u}_1 + \mathrm{u}_2) &= 0_V, \quad & T(r\mathrm{u}_1) &= 0_V \end{aligned}$$

since $\mathrm{u}_1 + \mathrm{u}_2, r\mathrm{u}_1 \in U$.

Hence we have

$$\begin{aligned} T(\mathrm{u}_1 + \mathrm{u}_2) &= 0_V = 0_V + 0_V = T(\mathrm{u}_1) + T(\mathrm{u}_2) \\ T(r\mathrm{u}_1) &= 0_V = r0_V = rT(\mathrm{u}_1). \end{aligned}$$

Since these equalities holds for all $\mathrm{u}_1, \mathrm{u}_2 \in U$, and $r \in \mathbb{F}$, the map $T : U \to V$ is a linear transformation.

(b) Determine the null space $\mathcal{N}(T)$ and the range $\mathcal{R}(T)$ of $T$.

The null space $\mathcal{N}(T)$ of $T$ is, by definition,

$$\mathcal{N}(T) = \{\mathrm{u} \in U \mid T(\mathrm{u}) = 0_V\}.$$

Since $T(\mathrm{u}) = 0_V$ for every $\mathrm{u} \in U$, we obtain

$$\mathcal{N}(T) = U.$$

The range $\mathcal{R}(T)$ of $T$ is, by definition,

$$\mathcal{R}(T) = \{\mathrm{v} \in V \mid \text{there exists } \mathrm{u} \in U \text{ such that } T(\mathrm{u}) = \mathrm{v}\}.$$

Since every vector of $U$ is mapped into $0_V$, we have

$$\mathcal{R}(T) = \{0_V\}.$$

7 Let $\mathbb{Q}$ denote the set of rational numbers (i.e., fractions of integers). Let $V$ denote the set of the form $x + y\sqrt{2}$ where $x, y \in \mathbb{Q}$. You may take for granted that the set $V$ is a vector space over the field $\mathbb{Q}$.
**(a)** Show that $B = \{1, \sqrt{2}\}$ is a basis for the vector space $V$ over $\mathbb{Q}$.
**(b)** Let $\alpha = a + b\sqrt{2} \in V$, and let $T_\alpha : V \to V$ be the map defined by

$$T_\alpha(x + y\sqrt{2}) := (ax + 2by) + (ay + bx)\sqrt{2} \in V$$

for any $x + y\sqrt{2} \in V$. Show that $T_\alpha$ is a linear transformation.
**(c)** Let $\begin{bmatrix} x \\ y \end{bmatrix}_B = x + y\sqrt{2}$. Find the matrix $T_B$ such that

$$T_\alpha(x + y\sqrt{2}) = \left(T_B \begin{bmatrix} x \\ y \end{bmatrix}\right)_B,$$

and compute $\det T_B$.
(*The Ohio State University*)

## Hint.

8 For (a), to show that $B$ is linearly independent, consider a linear combination

$$c_1 \cdot 1 + c_2\sqrt{2} = 0,$$

where $c_1, c_2 \in Q$. Use the fact that $\sqrt{2}$ is a irrational number to conclude $c_1 = c_2 = 0$.

9 You may directly check the linearity, or note that

$$T_\alpha(x + y\sqrt{2}) = \alpha(x + y\sqrt{2}).$$

Proof.

(a) Show that $B = \{1, \sqrt{2}\}$ is a basis for the vector space $V$ over $\mathbb{Q}$.

Since $V = \{x \cdot 1 + y\sqrt{2} \mid x, y \in \mathbb{Q}\} = \operatorname{Span}\{x, y\}$ , we see that $B = \{1, \sqrt{2}\}$ spans the vector space $V$.
To show that $B$ is a linearly independent set, consider a linear combination

$$c_1 \cdot 1 + c_2\sqrt{2} = 0,$$

where $c_1, c_2 \in \mathbb{Q}$.

---

We want to show that $c_1 = c_2 = 0$. Assume that this is not the case. Then both $c_1$ and $c_2$ must be nonzero.
Then we have

$$\sqrt{2} = -\frac{c_1}{c_2}.$$

(Note that $c_2 \neq 0$.)

Since $c_1, c_2 \in \mathbb{Q}$, the fraction $\frac{c_1}{c_2}$ is also a rational number. However $\sqrt{2}$ is irrational number.

Hence this is a contradiction and we must have $c_1 = c_2 = 0$. Thus $B$ is a linearly independent set, and $B$ is a basis.

(b) Show that $T_\alpha$ is a linear transformaiton.

First note that we can simplify the given formula as

$$T_\alpha(x + y\sqrt{2}) = (a + b\sqrt{2})(x + y\sqrt{2}) = \alpha(x + y\sqrt{2}).$$

Namely, $T_\alpha$ is a multiplication by $\alpha$.

Let $x_1 + y_1\sqrt{2}, x_2 + y_2\sqrt{2} \in V$, where $x_1, y_1, x_2, y_2 \in \mathbb{Q}$.
Then we have

$$\begin{aligned} T_\alpha\left((x_1 + y_1\sqrt{2}) + (x_2 + y_2\sqrt{2})\right) &= \alpha\left((x_1 + y_1\sqrt{2}) + (x_2 + y_2\sqrt{2})\right) \\ &= \alpha(x_1 + y_1\sqrt{2}) + \alpha(x_2 + y_2\sqrt{2}) \\ &= T_\alpha(x_1 + y_1\sqrt{2}) + T_\alpha(x_2 + y_2\sqrt{2}). \end{aligned}$$

For $x + y\sqrt{2} \in V$ and $c \in \mathbb{Q}$, we also have

$$\begin{aligned} T_\alpha\left(c(x + y\sqrt{2})\right) &= \alpha\left(c(x + y\sqrt{2})\right) \\ &= c\alpha(x + y\sqrt{2}) = cT_\alpha(x + y\sqrt{2}). \end{aligned}$$

Therefore $T_\alpha$ is a linear transformation.

(c) Find the matrix $T_B$ and compute $\det T_B$.

To find the matrix representation for the linear transformation $T_\alpha$ with respect to the basis $B$, we compute $T_\alpha(1)$ and $T_\alpha(\sqrt{2})$.

We have

$$\begin{aligned} T_\alpha(1) &= \alpha \cdot 1 = a + b\sqrt{2} = \begin{bmatrix} a \\ b \end{bmatrix}_B \\ T_\alpha(\sqrt{2}) &= \alpha \cdot \sqrt{2} = 2b + a\sqrt{2} = \begin{bmatrix} 2b \\ a \end{bmatrix}_B . \end{aligned}$$

Thus the matrix is

$$T_B = \begin{bmatrix} a & 2b \\ b & a \end{bmatrix}.$$

The determinant of this matrix is

$$\det T_B = a^2 - 2b^2.$$

### Comment.

In a typical first year linear algebra course, we use real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$ as a base field. Or just state theorems using a field $K$, for a general field $K$ but most of the examples use $\mathbb{R}$ or $\mathbb{C}$.

In this sense this problem is unorthodox because the field is rational numbers $\mathbb{Q}$.
But the problem is quite typical in Field theory or in Number theory.

⑩ Let $V$ be the vector space of all $2 \times 2$ real matrices and let $P_3$ be the vector space of all polynomials of degree $3$ or less with real coefficients. Let $T : P_3 \to V$ be the linear transformation defined by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{bmatrix} a_0 + a_2 & -a_0 + a_3 \\ a_1 - a_2 & -a_1 - a_3 \end{bmatrix}$$

for any polynomial $a_0 + a_1x + a_2x^2 + a_3 \in P_3$. Find a basis for the range of $T$, $\mathcal{R}(T)$, and determine the rank of $T$, $\operatorname{rank}(T)$, and the nullity of $T$, $\operatorname{null}(T)$. Also, prove that $T$ is not injective.

### Proof.

Let $\{1, x, x^2, x^3\}$ be a basis of the vector space $P_3$.
Then the range of $T$ is spanned by the images of these basis vectors. That is,

$$\mathcal{R}(T) = \operatorname{Span}\{T(1), T(x), T(x^2), T(x^3)\}.$$

Thus, the range $\mathcal{R}(T)$ is spanned by the following matrices

$$T(1) = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, T(x) = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$$

$$T(x^2) = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, T(x^3) = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}.$$

We find a basis among these matrices.

Let $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ be the standard basis of $V$, where the $(i, j)$-entry of $E_{ij}$ is 1 and the other entries are 0.

With respect to this basis $B$, the coordinate vectors of the above matrices are

$$[T(1)]_B = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, [T(x)]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

$$[T(x^2)]_B = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, [T(x^3)]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$

---

Let $A$ be the matrix whose columns are these vectors:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix}.$$

Applying the elementary row operations, we obtain

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} \xrightarrow{\substack{R_2+R_1 \\ R_4+R_3}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \end{bmatrix} \xrightarrow{\substack{R_1-R_3 \\ R_2+R_3 \\ R_4+R_3}} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is in reduced row echelon form and the first three columns contains the leading 1's.
Thus, the first three matrices $T(1)$, $T(x)$, $T(x^2)$ form a basis for the range $\mathcal{R}(T)$.
(This is the "leading 1 method".)

(Note also that the last column gives the coefficients of the linear combination
$T(x^3) = -T(1) + T(x) + T(x^2)$.)

---

It follows that the rank, which is the dimension of the range, is 3.

By the rank-nullity theorem we have

$$\operatorname{rank}(T) + \operatorname{null}(T) = \dim(P_3) = 4.$$

We deduce that the nullity is 1.

Finally, since the nullity of $T$ is 1, the linear transformation $T$ is not injective.
(Recall that $T$ is injective if and only if the nullity is 0.)

Another way to see that $T$ is not injective is as follows.
We saw that $T(x^3) = -T(1) + T(x) + T(x^2)$. This implies by linearity that

$$T(x^3) = T(-1 + x + x^2).$$

Thus the values of $T$ at two distinct polynomials $x^3$ and $-1 + x + x^2$ are the same.
Hence $T$ is not injective.

---

In conclusion,

$$\{T(1), T(x), T(x^2)\} = \left\{ \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \right\}$$

is a basis for the range of $T$. We have

$$\operatorname{rank}(T) = 3 \text{ and } \operatorname{null}(T) = 1.$$

---

11 Let $V$ be a real vector space of all real sequences

$$(a_i)_{i=1}^{\infty} = (a_1, a_2, \dots).$$

Let $U$ be the subspace of $V$ consisting of all real sequences that satisfy the linear recurrence relation $a_{k+2} - 5a_{k+1} + 3a_k = 0$ for $k = 1, 2, \dots$.
**(a)** Let

$$\begin{aligned} \mathbf{u}_1 &= (1, 0, -3, -15, -66, \dots) \\ \mathbf{u}_2 &= (0, 1, 5, 22, 95, \dots) \end{aligned}$$

be vectors in $U$. Prove that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis of $U$ and conclude that the dimension of $U$ is $2$.
**(b)** Let $T$ be a map from $U$ to $U$ defined by

$$T\big((a_1, a_2, \dots)\big) = (a_2, a_3, \dots).$$

Verify that the map $T$ actually sends a vector $(a_i)_{i=1}^{\infty} \in V$ to a vector $T\big((a_i)_{i=1}^{\infty}\big)$ in $U$, and show that $T$ is a linear transformation from $U$ to $U$.
**(c)** With respect to the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ obtained in (a), find the matrix representation $A$ of the linear transformation $T : U \to U$ from (b).

---

## Proof.

See the post Sequences satisfying linear recurrence relation form a subspace for a proof that $U$ is a subspace of $V$.

### (a) Prove that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis of $U$ and the dimension of $U$ is $2$.

Let

$$(a_i)_{i=1}^{\infty} = (a_1, a_2, a_3, \dots)$$

be an arbitrary vector in $U$.
If we know the first two terms $a_1, a_2$, the remaining terms are determined by the linear recurrence relation $a_{k+2} - 5a_{k+1} + 3a_k = 0$. Thus if the first two terms of two sequences of $U$ are equal, then the sequences are equal.
From this observation, we have

$$(a_i)_{i=1}^{\infty} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2.$$

Hence $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a spanning set of $U$.

---

It remains to show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a linearly independent set.
Suppose that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 = \mathbf{0},$$

where $\mathbf{0}$ is the zero sequence.
Then since we have

$$\begin{aligned}&(0, 0, 0 \dots)\\ &= c_1\mathbf{u}_1 + c_2\mathbf{u}_2\\ &= c_1(1, 0, -3, \dots) + c_2(0, 1, 5, \dots)\\ &= (c_1, c_2, -3c_1 + 5c_2, \dots),\end{aligned}$$

we must have $c_1 = c_2 = 0$. Hence $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a linearly independent set, and it is a basis. Therefore the dimension of $U$ is $2$.

> (b) $T$ is a linear transformation from $U$ to $U$.

Let us write $T\left((a_i)_{i=1}^{\infty}\right) = (b_i)_{i=1}^{\infty}$. Thus $b_i = a_{i+1}$ by definition of $T$. Then we have

$$b_{k+2} - 5b_{k+1} + 3b_k = a_{k+3} - 5a_{k+2} + 3a_{k+1} = 0$$

for $k = 1, 2, \dots$ since $(a_i)_{i=1}^{\infty} \in U$, hence satisfies the linear recurrence relation.
Thus the sequence $(b_i)_{i=1}^{\infty}$ also satisfies the recurrence relation, and it is in $U$. Thus $T$ sends a vector in $U$ to a vector in $U$.

---

Let us prove that $T$ is a linear transformation from $U$ to $U$.
Let $(a_i)_{i=1}^{\infty}, (b_i)_{i=1}^{\infty}$ be vectors in $U$ and let $r \in \mathbb{R}$ be a scalar.
Then we have

$$\begin{aligned}
&T\left((a_i)_{i=1}^{\infty}+(b_i)_{i=1}^{\infty}\right)\\
&=T\left((a_i+b_i)_{i=1}^{\infty}\right)\\
&=(a_2+b_2,a_3+b_3,\ldots)\\
&=(a_2,a_3,\ldots)+(b_2,b_3,\ldots)\\
&=T\left((a_i)_{i=1}^{\infty}\right)+T\left((b_i)_{i=1}^{\infty}\right)
\end{aligned}$$

Also we have

$$\begin{aligned}
&T\left(c(a_i)_{i=1}^{\infty}\right)\\
&=T\left((ca_i)_{i=1}^{\infty}\right)\\
&=(ca_2,ca_3,\ldots)\\
&=c(a_2,a_3,\ldots)\\
&=cT\left((a_i)_{i=1}^{\infty}\right)
\end{aligned}$$

Therefore $T$ is a linear transformation.

### (c) The matrix representation $A$ of the linear transformation $T$

We compute the values of the linear transformation of $T$ on basis vectors $\mathbf{u}_1,\mathbf{u}_2$ and express them as a linear combination of $\mathbf{u}_1,\mathbf{u}_2$.

We have

$$\begin{aligned}
T(\mathbf{u}_1)&=T\left((1,0,-3,-15,-66,\ldots)\right)\\
&=(0,-3,-15,-66,\ldots)\\
&=-3(0,1,5,22,\ldots)\\
&=0\mathbf{u}_1-3\mathbf{u}_2.
\end{aligned}$$

and also we have

$$\begin{aligned}
T(\mathbf{u}_2)&=T\left((0,1,5,22,95,\ldots)\right)\\
&=(1,5,22,95,\ldots)\\
&=(1,0,-3,-15,\ldots)+(0,5,25,110,\ldots)\\
&=\mathbf{u}_1+5\mathbf{u}_2.
\end{aligned}$$

Therefore the matrix representation of $T$ with respect to the basis $B=\{\mathbf{u}_1,\mathbf{u}_2\}$ is given by

$$A=[[T(\mathbf{u}_1)]_B,[T(\mathbf{u}_2)]_B]=\begin{bmatrix}0&1\\-3&5\end{bmatrix}.$$

Here $[T(\mathbf{u}_1)]_B=\begin{bmatrix}0\\-3\end{bmatrix}$ is the coordinate vector of $T(\mathbf{u}_1)$ with respect to the basis $B$. Similarly for $[T(\mathbf{u}_2)]_B]=\begin{bmatrix}1\\5\end{bmatrix}$.

Remark. Using the matrix $A$ for the linear transformation $T$, we can write formally

$$[T(\mathbf{u}_1)\quad T(\mathbf{u}_2)]=[\mathbf{u}_1\quad \mathbf{u}_2]\begin{bmatrix}0&1\\-3&5\end{bmatrix}.$$

**12** Let $V$ be the vector space of all $n\times n$ real matrices. Let us fix a matrix $A\in V$. Define a map $T:V\to V$ by $T(X)=AX-XA$ for each $X\in V$.

**(a)** Prove that $T:V\to V$ is a linear transformation.

**(b)** Let $B$ be a basis of $V$. Let $P$ be the matrix representation of $T$ with respect to $B$. Find the determinant of $P$.

## Proof.

### (a) Prove that $T : V \to V$ is a linear transformation.

To prove $T$ is a linear transformation, we need to show the following properties.

⑬ For any $X, Y \in V$, we have $T(X+Y) = T(X) + T(Y)$.

⑭ For any $X \in V, r \in \mathbb{R}$, we have $T(rX) = rT(X)$.

To check condition 1, let $X, Y \in V$. Then we have

$$\begin{aligned} T(X+Y) &= A(X+Y) - (X+Y)A && \text{by definition of } T \\ &= AX + AY - XA - YA \\ &= AX - XA + AY - YA \\ &= T(X) + T(Y) && \text{by definition of } T. \end{aligned}$$

Hence condition 1 is met.

To verify condition 2, let $X \in V, r \in \mathbb{R}$.

Then we have

$$\begin{aligned} T(rX) &= A(rX) - (rX)A && \text{by definition of } T \\ &= rAX - rXA && r \text{ is a scalar} \\ &= r(AX - XA) \\ &= rT(X) && \text{by definition of } T. \end{aligned}$$

So condition 2 is also met, hence $T$ is a linear transformation.

### (b) Find the determinant of the matrix representation of $T$.

Let $B$ be a basis of the vector space $V$ and let $P$ be the matrix of the linear transformation $T$ with respect to $B$. We prove that the determinant of $P$ is zero.

Let $I$ be the $n \times n$ identity matrix. Then we have

$$T(I) = AI - IA = A - A = O,$$

where $O$ is the $n \times n$ zero matrix.

Since $T(O) = O$, this implies that the linear transformation $T$ is not injective, hence $P$ is a singular matrix.

Let us explain the details.

Let $v = [I]_B \in \mathbb{R}^{n^2}$ be the coordinate vector of $I$ with respect to the basis $B$.

Then since $I \neq O$, the vector $v$ is not zero.

Then $T(I) = O$ implies that

$$Pv = 0 \in \mathbb{R}^{n^2}.$$

As the nonzero vector $v$ is a solution of the matrix equation $Px = 0$, the matrix $P$ is singular.

Since $P$ is singular, the determinant of $P$ is zero.

**15** The space $C^\infty(\mathbb{R})$ is the vector space of real functions which are infinitely differentiable. Let $T : C^\infty(\mathbb{R}) \to \mathrm{P}_3$ be the map which takes $f \in C^\infty(\mathbb{R})$ to its third order Taylor polynomial, specifically defined by

$$T(f)(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{6}x^3.$$

Here, $f'$, $f''$ and $f'''$ denote the first, second, and third derivatives of $f$, respectively. Prove that $T$ is a linear transformation.

**Proof.**

To prove that $T$ is a linear transformation, we must check that it satisfies the two axioms for linear transformations. First, suppose $f, g \in C^\infty(R)$. Then basic properties of the derivative tell us that

$$\begin{aligned}
(f+g)'(x) &= f'(x) + g'(x) \\
(f+g)''(x) &= f''(x) + g''(x) \\
(f+g)'''(x) &= f'''(x) + g'''(x).
\end{aligned}$$

And so we can see

$$\begin{aligned}
&T(f+g)(x) \\
&= (f+g)(0) + (f+g)'(0)x + \frac{(f+g)''(0)}{2}x^2 + \frac{(f+g)'''(0)}{6}x^3 \\
&= f(0) + g(0) + f'(0)x + g'(0)x + \frac{f''(0)}{2}x^2 + \frac{g''(0)}{2}x^2 + \frac{f'''(0)}{6}x^3 + \frac{g'''(0)}{6}x^3 \\
&= f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{6}x^3 + g(0) + g'(0)x + \frac{g''(0)}{2}x^2 + \frac{g'''(0)}{6}x^3 \\
&= T(f)(x) + T(g)(x).
\end{aligned}$$

---

Now for a scalar $c \in R$, the equality $(cf)'(x) = cf'(x)$ is a basic property of the derivative. Then,

$$\begin{aligned}
T(cf)(x) &= cf(0) + (cf)'(0)x + \frac{(cf)''(0)}{2}x^2 + \frac{(cf)'''(0)}{6}x^3 \\
&= cf(0) + cf'(0)x + c\frac{f''(0)}{2}x^2 + c\frac{f'''(0)}{6}x^3 \\
&= c\left(f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{6}x^3\right) \\
&= cT(f)(x)
\end{aligned}$$

Thus, $T$ is a linear transformation.

16 Let $P_n$ be the vector space of all polynomials with real coefficients of degree $n$ or less. Consider the differentiation linear transformation $T: P_n \to P_n$ defined by

$$T\,(\,f(x)\,)=\frac{d}{dx}f(x).$$

**(a)** Consider the case $n=2$. Let $B=\{1, x, x^2\}$ be a basis of $P_2$. Find the matrix representation $A$ of the linear transformation $T$ with respect to the basis $B$.
**(b)** Compute $A^3$, where $A$ is the matrix obtained in part (a).
**(c)** If you computed $A^3$ in part (b) directly, then is there any theoretical explanation of your result?
**(d)** Now we consider the general case. Let $B$ be any basis of the vector space of $P_n$ and let $A$ be the matrix representation of the linear transformation $T$ with respect to the basis $B$. Prove that without any calculation that the matrix $A$ is nilpotent.

## Background.

### Differentiating is a linear transformation.

Recall that differentiating a polynomial is a linear transformation.

See the post "Differentiation is a linear transformation" for a proof.

The post also contains a similar problem for $n=3$.

### Definition (Nilpotent matrix).

A square matrix $A$ is called **nilpotent** if there exists a positive integer $k$ such that $A^k=O$, the zero matrix.

## Solution.

### (a) Find the matrix representation $A$ of the linear transformation $T$ with respect to the basis $B$.

To find the matrix representation $A$ of $T$, we first find the coordinate vectors as follows.

Since we have

$$\begin{aligned} T(1) &= 0 = 0\cdot 1 + 0\cdot x + 0\cdot x^2 \\ T(x) &= 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ T(x^2) &= 2x = 0\cdot 1 + 2\cdot x + 0\cdot x^2 \end{aligned}$$

we have the coordinate vectors of these with respect to the basis $B$ is

$$[T(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(x)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(x^2)]_B = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

Hence the matrix representation $A$ of $T$ is

$$A = [\,[T(1)]_B \quad [T(x)]_B \quad [T(x^2)]_B\,] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

### (b) Compute $A^3$, where $A$ is the matrix obtained in part (a).

We calculate

$$\begin{aligned} A^2 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and },\\ A^3 = A^2 A &= \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Thus, $A^3$ is the zero matrix.

### (c) Is there any theoretical explanation of your result?

Recall from calculus that differentiating a polynomial once lowers the degree by one. Thus if we differentiate a degree $2$ polynomial once, then you get a degree $1$ polynomial. If we apply the differentiation again, we obtain a degree $0$ polynomial, which is a constant. If we differentiate once more, it becomes the zero polynomial.

Hence differentiating a polynomial of degree $2$ or less three times, the polynomial becomes the zero polynomial.

Differentiating once is the linear transformation $T$, hence multiplying by $A$ corresponds to differentiating once.
Therefore, $A^3$ corresponds to differentiating three times.
This is why $A^3$ is the zero matirx.

### (d) Prove that without any calculation that the matrix $A$ is nilpotent.

We generalize the argument in part (c) to any positive integer $n$. Recall again differentiating a polynomial once lowers the degree by one.
Thus, if we differentiate a degree $n$ polynomial $n+1$ times, then the polynomial becomes the zero polynomial.

In other words, we have

$$T^{n+1}(f(x))=0$$

for any $f(x)\in P_n$.

Thus we have $T^{n+1}$ is the zero linear transformation.

In terms of the matrix representation, this means that we have $A^{n+1}=O$, the zero matrix.

⑰ Let $W=C^\infty(\mathbb{R})$ be the vector space of all $C^\infty$ real-valued functions (smooth function, differentiable for all degrees of differentiation). Let $V$ be the vector space of all linear transformations from $W$ to $W$. The addition and the scalar multiplication of $V$ are given by those of linear transformations. Let $T_1, T_2, T_3$ be the elements in $V$ defined by

$$T_1(f(x))=\frac{\mathrm{d}}{\mathrm{d}x}f(x)$$

$$T_2(f(x))=\frac{\mathrm{d}^2}{\mathrm{d}x^2}f(x)$$

$$T_3(f(x))=\int_0^x f(t)\,\mathrm{d}t.$$

Then determine whether the set $\{T_1, T_2, T_3\}$ are linearly independent or linearly dependent.

## Proof.

We prove that the set $\{T_1, T_2, T_3\}$ are linearly independent.

Suppose that we have a linear combination

$$c_1T_1+c_2T_2+c_3T_3=0, \tag{*}$$

for some real numbers $c_1, c_2, c_3$.

Note that $0$ is the zero linear transformation $0(f(x))=0$ for all $x$, and it is the zero vector in the vector space $V$.

The equality (*) is an equality as linear transformations.

We need to show that $c_1=c_2=c_3=0$.

---

Let us consider the function $f(x)=1$ in $C^\infty(\mathbb{R})$.

Then we have

$$T_1(1)=0$$
$$T_2(1)=0$$
$$T_3(1)=\int_0^x 1\,\mathrm{d}t=x.$$

These yield that

$$0=c_1T_1(1)+c_2T_2(1)+c_3T_3(1)=c_3x,$$

and hence $c_3=0$.

Next, consider the function $f(x) = x$ .
We have

$$T_1(x) = 1$$
$$T_2(x) = 0$$

and

$$0 = c_1 T_1(x) + c_2 T_2(x) = c_1.$$

Thus $c_1 = 0$.

---

Finally, consider the function $f(x) = x^2$ .
We compute

$$T_2\left(x^2\right) = 2,$$

and the equality (*) together with $c_1 = c_3 = 0$ gives

$$0 = c_2 T_2\left(x^2\right) = 2c_2.$$

So $c_2 = 0$, and we have proved that $c_1 = c_2 = c_3 = 0$.
Thus we conclude that the set $\{T_1, T_2, T_3\}$ is linearly independent.

**18** Let $V$ denote the vector space of all real $2 \times 2$ matrices. Suppose that the linear transformation from $V$ to $V$ is given as below.

$$T(A) = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} A - A \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}.$$

Prove or disprove that the linear transformation $T : V \to V$ is an isomorphism.

## Hint.

A linear transformation $T$ is an **isomorphism** if it is both surjective (onto) and injective (one to one).

Recall that a linear transformation $T$ is injective if and only if the kernel is trivial, that is, $\ker(T) = \{0\}$.

## Proof.

**We claim that $T$ is not an isomorphism.**

Recall that an isomorphism means that $T$ is surjective (onto) and injective (one to one).
Thus, to disprove that $T$ is a linear transformation, it suffices to prove that $T$ is not surjective **or** $T$ is not injective.

Let us prove that $T$ is not injective.

Equivalently, we show that the kernel $\ker(T)$ is nontrivial, that is, $\ker(T) \neq \{O\}$, where $O$ is the $2 \times 2$ zero matrix.

If $A$ is in $\ker(T)$, then we have $O = T(A)$ and it follows that

$$\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} A = A \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}.$$

In other words, a matrix is in the kernel if it commutes with the matrix $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$.

---

So out goal is to find nonzero matrices that commute with $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$.

For example, the $2 \times 2$ identity matrix $I$ will do.

Hence the kernel of $T$ contains a nonzero matrix, hence $T$ is not injective.

Thus, $T$ is not isomorphism.

### Comments

Note that we did not have to determine the kernel.

What we needed is to show that $\ker(T) \neq \{O\}$.

We proved this by showing that $I \in \ker(T)$.

Another matrix in the kernel is $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$ itself.

Also, note that we proved that $T$ is not an isomorphism because $T$ is not injective.

We could have tried to prove that $T$ is not surjective but this is harder.

⑲ Let $U$ and $V$ be vector spaces over a scalar field $\mathbb{F}$. Let $T : U \to V$ be a linear transformation. Prove that $T$ is injective (one-to-one) if and only if the nullity of $T$ is zero.

### Definition (Injective, One-to-One Linear Transformation).

A linear transformation is said to be **injective** or **one-to-one** if provided that for all $\mathrm{u}_1$ and $\mathrm{u}_1$ in $U$, whenever $T(\mathrm{u}_1) = T(\mathrm{u}_2)$, then we have $\mathrm{u}_1 = \mathrm{u}_2$.

### Proof.

### $(\implies)$: If $T$ is injective, then the nullity is zero.

Suppose that $T$ is injective.

Our objective is to show that the null space $\mathcal{N}(T) = \{0_U\}$.

Since $T$ is a linear transformation, it sends the zero vector $0_U$ of $U$ to the zero vector $0_V$ of $V$.

In fact, we have

$$\begin{aligned} T(0_U) &= T(0_U - 0_U) \\ &= T(0_U + (-1)0_U) \\ &= T(0_U) + (-1)T(0_U) && \text{by linearity of } T \\ &= T(0_U) - T(0_U) = 0_V. \end{aligned}$$

Hence $0_U \in \mathcal{N}(T)$.

On the other hand, if $\mathbf{u} \in \mathcal{N}(T)$, then we have

$$T(\mathbf{u}) = 0_V = T(0_U).$$

Since $T$ is injective, it yields that $\mathbf{u} = 0_U$.
Therefore we obtain $\mathcal{N}(T) = \{0_U\}$, and the nullity of $T$ is zero.
(Recall that the nullity of $T$ is the dimension of $\mathcal{N}(T)$.)

### ( $\Longleftarrow$ ): If the nullity is zero, then $T$ is injective.

Next, suppose that the nullity of $T$ is zero.
This is equivalent to the condition $\mathcal{N}(T) = \{0_U\}$.
Our goal is to show that $T : U \to V$ is injective.

Suppose that $T(\mathbf{u}_1) = T(\mathbf{u}_2)$ for some $\mathbf{u}_1, \mathbf{u}_2 \in U$.
Then we have

$$\begin{aligned} 0_V &= T(\mathbf{u}_1) - T(\mathbf{u}_2) \\ &= T(\mathbf{u}_1) + (-1)T(\mathbf{u}_2) \\ &= T(\mathbf{u}_1 + (-1)\mathbf{u}_2) && \text{by linearity of } T \\ &= T(\mathbf{u}_1 - \mathbf{u}_2) \end{aligned}$$

It follows that the vector $\mathbf{u}_1 - \mathbf{u}_2$ is in the null space $\mathcal{N}(T) = \{0_U\}$.
Thus, we have $\mathbf{u}_1 - \mathbf{u}_2 = 0_U$, or $\mathbf{u}_1 = \mathbf{u}_2$.
So the linear transformation $T$ is injective.

## Related Question.

**Problem**.
Let $U$ and $V$ be finite dimensional vector spaces over a scalar field $\mathbb{F}$.
Consider a linear transformation $T : U \to V$.

Prove that if $\dim(U) > \dim(V)$, then $T$ cannot be injective.

The proof of this problem is given in the post ↴
A Linear Transformation $T : U \to V$ cannot be Injective if $\dim(U) > \dim(V)$

(20) Let $U$ and $V$ be finite dimensional vector spaces over a scalar field $\mathbb{F}$. Consider a linear transformation $T : U \to V$. Prove that if $\dim(U) > \dim(V)$, then $T$ cannot be injective (one-to-one).

## Hints.

You may use the folowing facts.

A linear transformation $T : U \to V$ is injective if and only if the nullity of $T$ is zero.

For the proof of this fact, see the post ↴
A Linear Transformation is Injective (One-To-One) if and only if the Nullity is Zero

**Rank-Nullity Theorem**
For a linear transformation $T : U \to V$, we have

$$\operatorname{null}(T) + \operatorname{rank}(T) = \dim(U).$$

---

We give two proofs. The first one uses the rank-nullity theorem.
The second one avoids using the theorem.

## Proof 1.

By the rank-nullity theorem, we have

$$\operatorname{null}(T) + \operatorname{rank}(T) = \dim(U).$$

Note that the rank of $T$ is the dimension of the range $\mathcal{R}(T)$, which is a subspace of the vector space of $V$.
Hence it yields that

$$\operatorname{rank}(T) = \dim(\mathcal{R}(T)) \leq \dim(V).$$

It follows that

$$\operatorname{null}(T) = \dim(U) - \operatorname{rank}(T) > \dim(U) - \dim(V) > 0,$$

where the last inequality follows from the assumption.

This implies that the nullity is nonzero, hence $T$ is not injective.
(See the hints above.)

## Proof 2.

In the second proof, we prove $T$ is injective without using the rank-nullity theorem.

Seeking a contradiction, assume that $T : U \to V$ is injective.
Let $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be a basis of the vector space $U$, where $n = \dim(U)$.
Put $\mathbf{v}_1 = T(\mathbf{u}_1), \dots, \mathbf{v}_n = T(\mathbf{u}_n)$.

We claim that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent in $V$.
To see this, consider the linear combination

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = 0_V,$$

where $c_1, \ldots, c_n$ are scalars in $\mathbb{F}$.
Then we have

$$\begin{aligned} 0_V &= c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \\ &= c_1T(\mathbf{u}_1) + \cdots + c_nT(\mathbf{u}_n) \\ &= T(c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n). \end{aligned}$$

Since $T$ is assumed to be injective, we must have

$$c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = 0_U.$$

Because $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are basis vectors, they are linearly independent.
It follows that

$$c_1 = \cdots = c_n = 0,$$

and hence the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.

---

As the vector space $V$ contains $n$ linearly independent vectors, we see that

$$\dim(V) \geq n = \dim(U),$$

which contradicts the assumption that $\dim(U) > \dim(V)$.
Therefore, $T$ cannot be injective.

21 Let $V$ be a vector space over the field of real numbers $\mathbb{R}$. Prove that if the dimension of $V$ is $n$, then $V$ is isomorphic to $\mathbb{R}^n$.

**Proof.**

Since $V$ is an $n$-dimensional vector space, it has a basis

$$B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\},$$

where each $\mathbf{v}_i$ is a vector in $V$.

---

Define a map $T: V \to \mathrm{R}^n$ by sending each vector $\mathbf{v} \in V$ to its coordinate vector $[\mathbf{v}]_B$ with respect to the basis $B$.
More explicitly, if

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \text{ with } c_1, \ldots, c_n \in \mathrm{R},$$

then the coordinate vector with respect to $B$ is

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in \mathrm{R}^n.$$

Then the map $T: V \rightarrow \mathrm{R}^n$ is defined by

$$T(\mathbf{v}) = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

---

It follows from the properties of the coordinate vectors that the map $T$ is a linear transformation.
We show that $T$ is bijective, hence an isomorphism.

### $T$ is injective.

To show that $T$ is injective, it suffices to show that the null space of $T$ is trivial: $\mathcal{N}(T) = \{\mathbf{0}\}$.
(See the post "A Linear Transformation is Injective (One-To-One) if and only if the Nullity is Zero" for a proof of this fact.)

If $\mathbf{v} \in \mathcal{N}(T)$, then we have

$$\mathbf{0} = T(\mathbf{v}) = [\mathbf{v}]_B.$$

So the coordinate vector of $\mathbf{v}$ is zero, hence we have

$$\mathbf{v} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n = \mathbf{0}.$$

Thus, $\mathcal{N}(T) = \{\mathbf{0}\}$, and $T$ is injective.

### $T$ is surjective.

To show that $T$ is surjective, let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

be an arbitrary vector in $\mathrm{R}^n$.
Then consider the vector

$$\mathbf{v} := a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

in $V$.

Then it follows from the definition of the linear transformation $T$ that

$$T(\mathbf{v}) = [\mathbf{v}]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \mathbf{a}.$$

Therefore $T$ is surjective.

In summary, $T: V \to \mathrm{R}^n$ is a bijective linear transformation, and hence $T$ is an isomorphism.

Thus, we conclude that $V$ and $\mathrm{R}^n$ are isomorphic.

## Generalization

We may use more general field F instead of the field of real numbers R.

Then the general statement is as follows.

If $V$ is an $n$-dimensional vector space over a field F, then $V$ is isomorphic to $\mathrm{F}^n$.

The proof is identical as above except that we replace R by F everywhere.

# Orthogonal Bases

**Definition**

Let $V$ be a subspace in $\mathbb{R}^n$.

1. If a basis $B$ for $V$ is an orthogonal set, then $B$ is called an **orthogonal basis**.
2. If a basis $B$ for $V$ is an orthonormal set, then $B$ is called an **orthonormal basis**.

**Summary**

Let $V$ be a subspace in $\mathbb{R}^n$.

1. From any basis $B$ of $V$, the Gram-Schmidt orthogonalization produces an orthogonal basis $B'$ for $V$.

## Problems

1. Let $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ be the set of the following vectors in $\mathbb{R}^4$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Find an orthogonal basis of the subspace $\mathrm{Span}(S)$ of $\mathbb{R}^4$.

### (a) Solution 1. (Using the Gram-Schmidt process)

It is straightforward to check that the vectors $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent, and hence the set $S$ is a basis of $\mathrm{Span}(S)$.

Since the dot (inner) product of $\mathbf{v}_1$ and $\mathbf{v}_2$ is

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 1 \neq 0,$$

$S$ is not an orthogonal basis. We apply the Gram-Schmidt process to generate an orthogonal basis from the basis $S$.

The Gram-Schmidt process for two vectors is as follows. We define vectors $\mathbf{u}_1, \mathbf{u}_2$ by the following formula. Then $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis of $\mathrm{Span}(S)$.

$$\begin{aligned} \mathrm{u}_1 &:= \mathrm{v}_1 \\ \mathrm{u}_2 &:= \mathrm{v}_2 - \frac{\mathrm{u}_1 \cdot \mathrm{v}_2}{\mathrm{u}_1 \cdot \mathrm{u}_1}\mathrm{u}_1. \end{aligned} \tag{*}$$

Since we have

$$\mathrm{u}_1 \cdot \mathrm{v}_2 = \mathrm{v}_1 \cdot \mathrm{v}_2 = 1 \text{ and } \mathrm{u}_1 \cdot \mathrm{u}_1 = \mathrm{v}_1 \cdot \mathrm{v}_1 = 2,$$

we compute

$$\begin{aligned} \mathrm{u}_2 &= \mathrm{v}_2 - \frac{1}{2}\mathrm{u}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix}. \end{aligned}$$

Therefore the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} \right\}$$

is an orthogonal basis of $\mathrm{Span}(S)$ .
Note that scaling by a nonzero scalar does not change the orthogonality, the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is also an orthogonal basis of $\mathrm{Span}(S)$ , just in case you prefer not to have a fraction.

### (a) Solution 2. (Using a pattern of the Gram-Schmidt process)

Here is another solution using a partial information of the Gram-Schmidt process.
As in Solution 1, the set $S$ is a (non-orthogonal) basis of $\mathrm{Span}(S)$ .
We want to apply the Gram-Schmidt process but suppose we only remember the pattern of the Gram-Schmidt process. Namely, we want to define orthogonal vectors $\mathrm{u}, \mathrm{u}_2$ by

$$\begin{aligned} \mathrm{u}_1 &:= \mathrm{v}_1 \\ \mathrm{u}_2 &:= \mathrm{v}_2 + a\mathrm{u}_1. \end{aligned}$$

Here $a$ is some number, which is given in the Gram-Schmidt process (*) but we don't remember.

We can still determine the number $a$ as follows. Since $\mathrm{u}_1$ and $\mathrm{u}_2$ will be orthogonal, we have

$$\begin{aligned} 0 &= \mathrm{u}_1 \cdot \mathrm{u}_2 = \mathrm{u}_1 \cdot (\mathrm{v}_2 + a\mathrm{u}_1) \\ &= \mathrm{u}_1 \cdot \mathrm{v}_2 + a\mathrm{u}_1 \cdot \mathrm{u}_1 \\ &= 1 + 2a. \end{aligned}$$

Hence, we obtain $a = -1/2$.

Then we determine

$$\begin{aligned} \mathrm{u}_2 &= \mathrm{v}_2 - \frac{1}{2}\mathrm{u}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix}. \end{aligned}$$

(So we could complete the Gram-Schmidt process even though we didn't remember the details.)

Hence the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ 1/2 \\ 0 \end{bmatrix} \right\}$$

is an orthogonal basis of $\mathrm{Span}(S)$ .

## (b) Solution.

We first express the vector $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ as the linear combination

$$\begin{bmatrix} 3 \\ -2 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3\mathrm{e}_1 - 2\mathrm{e}_2.$$

Then we compute

$$\begin{aligned} T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) &= T(3\mathrm{e}_1 - 2\mathrm{e}_2) \\ &= 3T(\mathrm{e}_1) - 2T(\mathrm{e}_2) && \text{by linearity of } T \\ &= 3\begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 8 \\ 2 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 15 \\ 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 16 \\ 4 \\ 12 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -6 \end{bmatrix}. \end{aligned}$$

Therefore we have found

$$T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -1 \\ -6 \end{bmatrix}.$$

② Let $W$ be a subspace of $\mathbb{R}^4$ with a basis

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Find an ortho**normal** basis of $W$.
*(The Ohio State University)*

## Solution.

Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Note that they are not orthogonal as the dot product is

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 2 \neq 0.$$

---

Let us first find an orthogonal basis for $W$ by the Gram-Schmidt orthogonalization process.

Let $\mathbf{w}_1 := \mathbf{v}_1$.
Next, let $\mathbf{w}_2 := \mathbf{v}_2 + a\mathbf{v}_1$, where $a$ is a scalar to be determined so that $\mathbf{w}_1 \cdot \mathbf{w}_2 = 0$.
(You may also use the formula of the Gram-Schmidt orthogonalization.)

As $\mathbf{w}_1$ and $\mathbf{w}_2$ is orthogonal, we have

$$\begin{aligned} 0 &= \mathbf{w}_1 \cdot \mathbf{w}_2 = \mathbf{v}_1 \cdot (\mathbf{v}_2 + a\mathbf{v}_1) \\ &= \mathbf{v}_1 \cdot \mathbf{v}_2 + a\mathbf{v}_1 \cdot \mathbf{v}_1 \\ &= 2 + 3a. \end{aligned}$$

It follows that $a = -2/3$ and

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{2}{3}\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

---

Now, to avoid fractions in our computation, let us consider $3\mathbf{w}_2$, instead of $\mathbf{w}_2$. Note that the scaling does not change the orthogonality.
We have

$$3\mathbf{w}_2 = 3\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ 1 \\ 1 \end{bmatrix}.$$

Thus the set $\{\mathbf{w}_1, 3\mathbf{w}_2\}$ is an orthogonal basis for $W$.
However, the length of these vectors are not 1 as we see

$$\begin{aligned} \|\mathbf{w}_1\| &= \sqrt{1^2 + 0^2 + 1^2 + 1^2} = \sqrt{3} \\ \|3\mathbf{w}_2\| &= \sqrt{(-2)^2 + 3^2 + 1^2 + 1^2} = \sqrt{15}. \end{aligned}$$

---

Now it suffices to normalize the vectors $\mathbf{w}_1, 3\mathbf{w}_2$ to obtain an orthonormal basis.
Therefore, the set

$$\left\{ \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{15}}\begin{bmatrix} -2 \\ 3 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

is an orthonormal basis for $W$.

## Comment.

This is one of the midterm 2 exam problems for Linear Algebra (Math 2568) in Autumn 2017.

One common mistake is just to normalize the vectors by dividing them by their length $\sqrt{3}$.
The resulting vectors have length 1, but they are not orthogonal.

Another mistake is that you just changed the numbers in the vectors so that they are orthogonal.
The issue here is that if you change the numbers randomly, then the new vectors might no longer

belong to the subspace $W$.

The point of the Gram-Schmidt orthogonalization is that the process converts any basis for $W$ to an orthogonal basis for $W$.

The above solution didn't use the full formula of the Gram-Schmidt orthogonalization. Of course, you may use the formula in the exam but you must remember it correctly.

---

③ Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

**(a)** Find an orthonormal basis of the null space of $A$.

**(b)** Find the rank of $A$.

**(c)** Find an orthonormal basis of the row space of $A$.

(*The Ohio State University*)

## Solution.

First of all, note that $A$ is already in reduced row echelon form.

### (a) Find an orthonormal basis of the null space of $A$.

Let us find a basis of null space of $A$.

The null space consists of the solutions of $A\mathbf{x} = 0$.

Since $A$ is in reduced row echelon form, the solutions $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ satisfy

$$x_1 = -x_3 \text{ and } x_2 = 0,$$

hence the general solution is

$$\mathbf{x} = \begin{bmatrix} -x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, the set

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis of the null space of $A$.

Since the length of the basis vector is $\sqrt{(-1)^2 + 0^2 + 1^2} = \sqrt{2}$, it is not orthonormal basis.

Thus, we divide the vector by its length and obtain an orthonormal basis

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

### (b) Find the rank of $A$.

From part (a), we see that the nullity of $A$ is 1. The rank-nullity theorem says that

$$\text{rank of } A + \text{ nullity of } A = 3.$$

Hence the rank of $A$ is 2.

The second way to find the rank of $A$ is to use the definition of the rank: The rank of a matrix $B$ is the number of nonzero rows in a reduced row echelon matrix that is row equivalent to $B$.
Since $A$ is in echelon form and it has two nonzero rows, the rank is 2.

The third way to find the rank is to use the leading 1 method. By the leading 1 method, we see that the first two columns form a basis of the range, hence the rank of $A$ is 2.

### (c) Find an orthonormal basis of the row space of $A$.

By the row space method, the nonzero rows in reduced row echelon form a basis of the row space of $A$. Thus

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is a basis of the row space of $A$.
Since the dot (inner) product of these two vectors is 0, they are orthogonal.
The length of the vectors is $\sqrt{2}$ and 1, respectively.
Hence an orthonormal basis of the row space of $A$ is

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

---

④ Let $\mathbf{v}_1 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ be a vector in $\mathbb{R}^3$. Find an orthonormal basis for $\mathbb{R}^3$ containing the vector $\mathbf{v}_1$.

The first solution uses the Gram-Schumidt orthogonalization process.
On the other hand, the second solution uses the cross product.

## Solution 1 (The Gram-Schumidt Orthogonalization)

First of all, note that the length of the vector $\mathbf{v}_1$ is 1 as

$$\|\mathbf{v}_1\| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = 1.$$

---

We want to find two vectors $\mathbf{v}_2, \mathbf{v}_3$ such that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis for $\mathrm{R}^3$.
The vectors $\mathbf{v}_2, \mathbf{v}_3$ must lie on the plane that is perpendicular to the vector $\mathbf{v}_1$.
So consider the subspace

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathrm{R}^3 \;\middle|\; \begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix} = 0 \right\}$$

Note that $W$ consists of all vectors that are perpendicular to $\mathbf{v}_1$, hence $W$ is a plane that is perpendicular to $\mathbf{v}_1$.

---

The relation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix} = 0$$

can be written as

$$\frac{2}{3}x + \frac{2}{3}y + \frac{1}{3}z = 0,$$

or equivalently

$$z = -2x - 2y.$$

Hence the vectors in $W$ can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ -2x - 2y \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

It follows that

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} \right\}$$

is a basis for the subspace $W$. Let us call these vectors $\mathbf{u}_1, \mathbf{u}_2$, respectively.

---

We apply the Gram-Schmidt orthogonalization to this basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ and obtain an orthogonal basis as follows.
We do not change the first vector: let $\mathbf{w}_1 = \mathbf{u}_1$.

Next, we set

$$\mathbf{w}_2 = \mathbf{u}_2 + a\mathbf{u}_1$$

for some scalar $a$.

To determine $a$, we compute

$$\begin{aligned} 0 = \mathbf{w}_1 \cdot \mathbf{w}_2 &= \mathbf{u}_1 \cdot (\mathbf{u}_2 + a\mathbf{u}_1) \\ &= \mathbf{u}_1 \cdot \mathbf{u}_2 + a\mathbf{u}_1 \cdot \mathbf{u}_1 \\ &= (1 \cdot 0 + 0 \cdot 1 + (-2) \cdot (-2)) + a(1 \cdot 1 + 0 \cdot 0 + (-2) \cdot (-2) \\ &= 4 + 5a. \end{aligned}$$

Hence, $a = -4/5$ and we obtain

$$\mathbf{w}_2 = \mathbf{u}_2 - \frac{4}{5}\mathbf{u}_1$$

$$= \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} - \frac{4}{5}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$$

(Note that you may use the Gram-Schumidt orthogonalization formula, instead of the above method.)

---

As scaling does not change the orthogonality, consider $5\mathbf{w}_2$, instead of $\mathbf{w}_2$ (to avoids fractions).

We have

$$5\mathbf{w}_2 = \begin{bmatrix} 0 \\ 5 \\ -10 \end{bmatrix} - 4\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -4 \\ 5 \\ -2 \end{bmatrix}.$$

Therefore, $\{\mathbf{w}_1, 5\mathbf{w}_2\}$ is an orthogonal basis of $W$.

---

We obtain an orthonormal basis of $W$ by normalizing the length of these basis vectors.

As

$$\|\mathbf{w}_1\| = \sqrt{1^2 + 0^2 + (-2)^2} = \sqrt{5}$$

and

$$\|5\mathbf{w}_2\| = \sqrt{(-4)^2 + 5^2 + (-2)^2} = \sqrt{45} = 3\sqrt{5},$$

the vectors

$$\mathbf{v}_2 := \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

and

$$\mathbf{v}_3 := \frac{5\mathbf{w}_2}{\|5\mathbf{w}_2\|} = \frac{1}{3\sqrt{5}}\begin{bmatrix} -4 \\ 5 \\ -2 \end{bmatrix}$$

form an orthonormal basis of $W$.

Note that as the vectors $\mathbf{v}_2, \mathbf{v}_3$ lie in $W$, they are still perpendicular to the vector $\mathbf{v}_1$.

---

It follows that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonomal set in $\mathrm{R}^3$, thus it is an orthonormal basis for $\mathrm{R}^3$.

## Solution 2 (Cross Product)

Next, we solve the problem using the cross product.

Let $\mathbf{v}_1 = \frac{1}{3}\mathbf{u}_1$, where $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$.

Our first goal is to find the vectors $\mathbf{u}_2$ and $\mathbf{u}_3$ such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathrm{R}^3$.

---

Let $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be a vector that is perpendicular to $\mathbf{u}_1$.

Then we have $\mathbf{x} \cdot \mathbf{u}_1 = 0$, and hence we have the relation

$$2x + 2y + z = 0.$$

For example, the vector $\mathbf{u}_2 := \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ satisfies the relation, and hence $\mathbf{u}_2 \cdot \mathbf{u}_1 = 0$.

(So far, it is not so different from Solution 1.)

---

Now, let us define the third vector $\mathbf{u}_3$ to be the cross product of $\mathbf{u}_1$ and $\mathbf{u}_2$:

$$\mathbf{u}_3 := \mathbf{u}_1 \times \mathbf{u}_2 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -4 \\ 5 \\ -2 \end{bmatrix}.$$

By the property of the cross product, the vector $\mathbf{u}_3$ is perpendicular to both $\mathbf{u}_1, \mathbf{u}_2$.

Therefore, the set

$$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -2 \end{bmatrix} \right\}$$

is an orthogonal basis for $R^3$ as it consists of three nonzero orthogonal vectors.

---

Finally, to obtain an orthonormal basis for $R^2$, we just need to normalize the lengths of these vectors. The lengths are

$$\begin{aligned} \|\mathbf{u}_1\| &= \sqrt{2^2 + 2^2 + 1^2} = \sqrt{9} = 3 \\ \|\mathbf{u}_2\| &= \sqrt{1^2 + 0^2 + (-2)^2} = \sqrt{5} \\ \|\mathbf{u}_3\| &= \sqrt{(-4)^2 + 5^2 + (-2)^2} = \sqrt{45} = 3\sqrt{5}. \end{aligned}$$

Then we have $\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$.

Let $\mathbf{v}_2 := \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$ and $\mathbf{v}_3 := \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}$.

It follows that the set

$$\{\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3\} = \left\{ \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \frac{1}{3\sqrt{5}}\begin{bmatrix} -4 \\ 5 \\ -2 \end{bmatrix} \right\}$$

is an orthonormal basis for $R^3$.

5 Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a set of nonzero vectors in $\mathbb{R}^n$. Suppose that $S$ is an orthogonal set.
**(a)** Show that $S$ is linearly independent.
**(b)** If $k = n$, then prove that $S$ is a basis for $\mathbb{R}^n$.

## Proof.

### (a) Show that $S$ is linearly independent.

Consider the linear combination

$$c_1 \mathrm{v}_1 + c_2 \mathrm{v}_2 + \cdots + c_k \mathrm{v}_k = 0.$$

Our goal is to show that $c_1 = c_2 = \cdots = c_k = 0$.

---

We compute the dot product of $\mathrm{v}_i$ and the above linear combination for each $i = 1, 2, \dots, k$:

$$\begin{aligned} 0 &= \mathrm{v}_i \cdot 0 \\ &= \mathrm{v}_i \cdot (c_1 \mathrm{v}_1 + c_2 \mathrm{v}_2 + \cdots + c_k \mathrm{v}_k) \\ &= c_1 \mathrm{v}_i \cdot \mathrm{v}_1 + c_2 \mathrm{v}_i \cdot \mathrm{v}_2 + \cdots + c_k \mathrm{v}_i \cdot \mathrm{v}_k. \end{aligned}$$

As $S$ is an orthogonal set, we have $\mathrm{v}_i \cdot \mathrm{v}_j = 0$ if $i \neq j$.

Hence all terms but the $i$-th one are zero, and thus we have

$$0 = c_i \mathbf{v}_i \cdot \mathbf{v}_i = c_i \|\mathbf{v}_i\|^2.$$

Since $\mathbf{v}_i$ is a nonzero vector, its length $\|\mathbf{v}_i\|$ is nonzero.
It follows that $c_i = 0$.

As this computation holds for every $i = 1, 2, \dots, k$, we conclude that $c_1 = c_2 = \cdots = c_k = 0$.
Hence the set $S$ is linearly independent.

(b) If $k = n$, then prove that $S$ is a basis for $\mathbb{R}^n$.

Suppose that $k = n$. Then by part (a), the set $S$ consists of $n$ linearly independent vectors in the dimension $n$ vector space $\mathbb{R}^n$.

Thus, $S$ is also a spanning set of $\mathbb{R}^n$, and hence $S$ is a basis for $\mathbb{R}^n$.

# Determinants of Matrices

**Summary**

Let $A, B$ be $n \times n$ matrices.

① $A$ is nonsingular if and only if $\det(A) \neq 0$.

② $\det(AB) = \det(A)\det(B)$.

③ If $A$ is invertible, then $\det(A^{-1}) = \det(A)^{-1}$.

## Problems

① Let $A$ and $B$ be $n \times n$ matrices, where $n$ is an integer greater than $1$. Is it true that

$$\det(A+B) = \det(A) + \det(B)?$$

If so, then give a proof. If not, then give a counterexample.

### Solution.

We claim that the statement is false.

As a counterexample, consider the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and we have

$$\det(A+B) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

On the other hand, the determinants of $A$ and $B$ are

$$\det(A) = 0 \text{ and } \det(B) = 0,$$

and hence

$$\det(A) + \det(B) = 0 \neq 1 = \det(A+B).$$

Therefore, the statement is false and in general we have

$$\det(A+B) \neq \det(A) + \det(B).$$

### Remark.

When we computed the $2 \times 2$ matrices, we used the formula

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

This problem showed that the determinant does not preserve the addition.

However, the determinant is multiplicative.

In general, the following is true:

$$\det(AB) = \det(A)\det(B).$$

---

② Let

$$A = \begin{bmatrix} 2 & 0 & 10 \\ 0 & 7+x & -3 \\ 0 & 4 & x \end{bmatrix}.$$

Find all values of $x$ such that $A$ is invertible.
(*Stanford University*)

### Hint.

Calculate the determinant of the matrix $A$.

### Solution.

A matrix is invertible if and only if its determinant is non-zero.

So we first calculate the determinant of the matrix $A$.

By the first column cofactor expansion, we have

$$\begin{aligned} \det(A) &= 2\begin{vmatrix} 7+x & -3 \\ 4 & x \end{vmatrix} \\ &= 2\left((7+x)x - (-3)4\right) = 2(x^2 + 7x + 12) \\ &= 2(x+3)(x+4). \end{aligned}$$

Thus the determinant of $A$ is zero if and only if $x = -3$ or $x = -4$.

Therefore the matrix $A$ is invertible for all $x$ except $x = -3$ and $x = -4$.

---

③ Let

$$A = \begin{bmatrix} 1 & -x & 0 & 0 \\ 0 & 1 & -x & 0 \\ 0 & 0 & 1 & -x \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

be a $4 \times 4$ matrix. Find all values of $x$ so that the matrix $A$ is singular.

### Hint.

Use the fact that a matrix is singular if and only if the determinant of the matrix is zero.

To compute the determinant, use a cofactor expansion.

### Solution.

We use the fact that a matrix is singular if and only if the determinant of the matrix is zero.

We compute the determinant of $A$ as follows.

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1 & -x & 0 & 0 \\ 0 & 1 & -x & 0 \\ 0 & 0 & 1 & -x \\ 0 & 1 & 0 & -1 \end{vmatrix} \\
&= \begin{vmatrix} 1 & -x & 0 \\ 0 & 1 & -x \\ 1 & 0 & -1 \end{vmatrix} - 0 \begin{vmatrix} -x & 0 & 0 \\ 0 & 1 & -x \\ 1 & 0 & -1 \end{vmatrix} + 0 \begin{vmatrix} -x & 0 & 0 \\ 1 & -x & 0 \\ 1 & 0 & -1 \end{vmatrix} + 0 \begin{vmatrix} -x & 0 & 0 \\ 1 & -x & 0 \\ 0 & 1 & -x \end{vmatrix} \\
&\text{by the first column cofactor expansion} \\
&= \begin{vmatrix} 1 & -x & 0 \\ 0 & 1 & -x \\ 1 & 0 & -1 \end{vmatrix} \\
&= \begin{vmatrix} 1 & -x \\ 0 & -1 \end{vmatrix} - 0 \begin{vmatrix} -x & 0 \\ 0 & -1 \end{vmatrix} + \begin{vmatrix} -x & 0 \\ 1 & -x \end{vmatrix} \\
&\text{by the first column cofactor expansion} \\
&= -1 + x^2.
\end{aligned}$$

Therefore we have $\det(A) = x^2 - 1$. Thus $\det(A) = 0$ if and only if $x = \pm 1$.
We conclude that the matrix $A$ is singular if and only if $x = \pm 1$.

### Comment.

You may use the rule of Sarrus to compute the $3 \times 3$ determinant instead of the cofactor expansion if you like so.

---

4 Find all the values of $x$ so that the following matrix $A$ is a singular matrix.

$$A = \begin{bmatrix} x & x^2 & 1 \\ 2 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix}.$$

### Hint.

Use the fact that a matrix is singular if and only if its determinant is zero.

### Solution.

Note that a matrix is singular if and only if its determinant is zero.
So we compute the determinant of the matrix $A$ as follows.

$$\begin{aligned}\det(A) &= \begin{vmatrix} x & x^2 & 1 \\ 2 & 3 & 1 \\ 0 & -1 & 1 \end{vmatrix} \\ &= (-1)^{3+1}\cdot 0 \cdot \begin{vmatrix} x^2 & 1 \\ 3 & 1 \end{vmatrix} + (-1)^{3+2}\cdot(-1)\cdot \begin{vmatrix} x & 1 \\ 2 & 1 \end{vmatrix} + (-1)^{3+3}\cdot 1 \cdot \begin{vmatrix} x & x^2 \\ 2 & 3 \end{vmatrix} \\ &\text{by the third row cofactor expansion} \\ &= 0 + (x-2) + (3x - 2x^2) \\ &= -2x^2 + 4x - 2.\end{aligned}$$

Thus the determinant of $A$ is zero if

$$\det(A) = -2x^2 + 4x - 2 = 0,$$

equivalently,

$$x^2 - 2x + 1 = (x-1)^2 = 0.$$

Thus, the determinant of the matrix $A$ is zero if and only if $x = 1$.
Hence the matrix $A$ is singular if and only if $x = 1$.

---

5 Find the value(s) of $h$ for which the following set of vectors

$$\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} h \\ 1 \\ -h \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2h \\ 3h+1 \end{bmatrix} \right\}$$

is linearly independent.
(*Boston College*)

### Solution 1.

Let us consider the linear combination

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{0}. \qquad (*)$$

If this homogeneous system has only zero solution $x_1 = x_2 = x_3 = 0$, then the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.
We reduce the augmented matrix for the system as follows.

$$\left[\begin{array}{ccc|c} 1 & h & 1 & 0 \\ 0 & 1 & 2h & 0 \\ 0 & -h & 3h+1 & 0 \end{array}\right] \xrightarrow{R_3 + hR_2} \left[\begin{array}{ccc|c} 1 & h & 1 & 0 \\ 0 & 1 & 2h & 0 \\ 0 & 0 & 2h^2+3h+1 & 0 \end{array}\right].$$

From this, we see that the homogeneous system (*) has only the zero solution if and only if

$$2h^2 + 3h + 1 \neq 0.$$

Since we have

$$2h^2 + 3h + 1 = (2h+1)(h+1),$$

if $h \neq -\frac{1}{2}, -1$, then $2h^2 + 3h + 1 \neq 0$.

In summary, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent for any $h$ except $-\frac{1}{2}, -1$.

### Solution 2.

Note that the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent if and only the matrix

$$A := [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3] = \begin{bmatrix} 1 & h & 1 \\ 0 & 1 & 2h \\ 0 & -h & 3h+1 \end{bmatrix}$$

is nonsingular.

Also, the matrix $A$ is nonsingular if and only if the determinant $\det(A)$ is nonzero.
So we compute the determinant of the matrix $A$ by the first column cofactor expansion and obtain

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & h & 1 \\ 0 & 1 & 2h \\ 0 & -h & 3h+1 \end{vmatrix} \\ &= \begin{vmatrix} 1 & 2h \\ -h & 3h+1 \end{vmatrix} \\ &= 2h^2 + 3h + 1 \\ &= (2h+1)(h+1). \end{aligned}$$

Hence $\det(A) \neq 0$ if and only if $h \neq -\frac{1}{2}, -1$.
Therefore, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent for any values of $h$ except $-\frac{1}{2}, -1$.

---

⑥ Let

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 4 & 5 & 6 \\ 7 & 0 & 9 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Then find the value of

$$\det(A^2 B^{-1} A^{-2} B^2).$$

(Without a proof, you may assume that $A$ and $B$ are invertible matrices.)

### (a) Solution.

To determine eigenvalues of $A$, we compute the determinant of $A - \lambda I$.
We have

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & -2 \\ 6 & -4-\lambda \end{vmatrix} \\ &= (3-\lambda)(-4-\lambda) + 12 \\ &= \lambda^2 + \lambda = \lambda(\lambda+1). \end{aligned}$$

The eigenvalues are solutions of $\det(A - \lambda I) = 0$, hence eigenvalues of $A$ are $0, -1$.

---

Next, we find the eigenvector corresponding to the eigenvalue $\lambda = 0$.
Eigenvectors $\mathbf{x}$ are nonzero solutions of $(A - 0I)\mathbf{x} = \mathbf{0}$.

Thus, we solve $A\mathbf{x} = \mathbf{0}$. The augmented matrix of the system is

$$\left[\begin{array}{cc|c} 3 & -2 & 0 \\ 6 & -4 & 0 \end{array}\right] \xrightarrow{R_2-2R_1} \left[\begin{array}{cc|c} 3 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{cc|c} 1 & -2/3 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Thus, if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a solution, then $x_1 = \frac{2}{3}x_2$, hence

$$\mathbf{x} = x_2 \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to $\lambda = 0$ for any nonzero scalar $x_2$.

---

Finally, we find the eigenvectors corresponding to the eigenvalue $\lambda = -1$.

In this case we need to solve $(A - (-1)I)\mathbf{x} = \mathbf{0}$.

The augmented matrix is

$$\left[\begin{array}{cc|c} 4 & -2 & 0 \\ 6 & -3 & 0 \end{array}\right] \xrightarrow[\frac{1}{3}R_2]{\frac{1}{2}R_1} \left[\begin{array}{cc|c} 2 & -1 & 0 \\ 2 & -1 & 0 \end{array}\right] \xrightarrow{R_2-R_1} \left[\begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{2}R_1} \left[\begin{array}{cc|c} 1 & -1/2 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

It follows that the eigenvectors associated to $\lambda = -1$ are

$$\mathbf{x} = x_2 \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

for any nonzero scalar $x_2$.

## (b) Solution.

We use the following two properties of determinants.

Let $C$ and $D$ be $n \times n$ matrices. Then we have

$$\det(CD) = \det(C)\det(D)$$

and if $C$ is invertible, then

$$\det(C^{-1}) = \det(C)^{-1} = \frac{1}{\det(C)}.$$

Using the properties of determinants, we compute

$$\begin{aligned} &\det(A^2B^{-1}A^{-2}B^2) \\ &= \det(A)^2\det(B)^{-1}\det(A)^{-2}\det(B)^2 \\ &= \det(A)^2\det(A)^{-2}\det(B)^{-1}\det(B)^2 \qquad \text{(determinants are just numbers)} \\ &= \det(B). \end{aligned}$$

Hence it suffices to find the determinant of the matrix $B$.

Since the matrix $B$ is an upper triangular matrix, its determinant is the product of diagonal entries, thus

$$\det(B) = 2 \cdot 3 \cdot 4 = 24.$$

As a result, we obtain

$$\det(A^2 B^{-1} A^{-2} B^2) = 24.$$

---

⑦ Let $A$ be an $n \times n$ nilpotent matrix, that is, $A^m = O$ for some positive integer $m$, where $O$ is the $n \times n$ zero matrix.
Prove that $A$ is a singular matrix and also prove that $I - A, I + A$ are both nonsingular matrices, where $I$ is the $n \times n$ identity matrix.

## Hint.

We give two proofs. The first one uses only some properties of the determinant of a matrix.

The second one uses the theory of eigenvalues.

We summarize several facts we will use below.

In proof 1.

⑧ A matrix is nonsingular if and only if it is invertible.

⑨ A matrix is nonsingular if and only if its determinant is nonzero.

⑩ For two square matrices $A$, $B$, we have

$$\det(AB) = \det(A)\det(B).$$

In proof 2.

⑪ Eigenvalues of $A$ are all scalars $\lambda$ such that the matrix $A - \lambda I$ is nonsingular.

## Proof 1.

We use the fact that a matrix is nonsingular if and only if its determinant is nonzero.

Since the determinant is multiplicative, we have

$$0 = \det(O) = \det(A^m) = \det(A)^m.$$

This implies that $\det(A) = 0$, and hence the matrix $A$ is singular.

Next, we show that the matrix $I-A$ is nonsingular. Note that we have

$$I=I-A^m=(I-A)(I+A+A^2+\cdots+A^{m-1}).$$

This implies that the matrix $I+A+A^2+\cdots+A^{m-1}$ is the inverse matrix of $I-A$.
Thus the matrix $I-A$ is invertible, and hence it is nonsingular.

---

Next, we show that the matrix $I+A$ is nonsingular.
Note that we have

$$I-(-A)^m=(I+A)(I-A+A^2-\cdots+(-1)^{m-1}A^{m-1}).$$

Since $(-A)^m=(-1)^mA^m=(-1)^mO=O$, we obtain

$$I=(I+A)(I-A+A^2-\cdots+(-1)^{m-1}A^{m-1})$$

and this implies that $I+A$ is invertible matrix with the inverse matrix

$$(I+A)^{-1}=I-A+A^2-\cdots+(-1)^{m-1}A^{m-1}.$$

Therefore, the matrix $I+A$ is nonsingular.

## Proof 2.

In the second proof, we will use an eigenvalue technique.

Let $\lambda$ be an eigenvalue of the matrix $A$ and let $\mathbf{x}$ be the corresponding eigenvector. Namely, we have

$$A\mathbf{x}=\lambda\mathbf{x}.$$

Then we have

$$\begin{aligned}\mathbf{0}&=O\mathbf{x}=A^m\mathbf{x}\\&=A^{m-1}A\mathbf{x}=A^{m-1}(\lambda\mathbf{x})=\cdots=\lambda^m\mathbf{x}.\end{aligned}$$

It follows that $\lambda=0$. Thus the only eigenvalue of $A$ is zero.

Recall that eigenvalues $\lambda$ are all scalars such that $A-\lambda I$ is a singular matrix.
Since $\lambda=0$ is an eigenvalue of $A$, the matrix $A=A-0I$ is singular.

Since 1 is not an eigenvalue, we see that $A-I$ is nonsingular, and hence $I-A=-(A-I)$ is also nonsingular.

Similarly, since $-1$ is not an eigenvalue, we observe that $I+A=A+I$ is nonsingular.

## Related Question.

A natural source of nilpotent matrices is the differentiation linear transformation of vector space of polynomials.

Check out the post "Differentiating Linear Transformation is Nilpotent" for problems and solutions about the nilpotency of the differentiation linear transformation.

## Related Question.

Here is another problem about a nilpotent matrix.

**Problem.**

Suppose that $A$ is an $n \times n$ nilpotent matrix and $B$ is an $n \times n$ invertible matrix.

Is the matrix $B - A$ invertible? If so, give a proof. Otherwise, give a counterexample.

---

⑫ Let $A$ be a $3 \times 3$ matrix. Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are linearly independent $3$-dimensional vectors. Suppose that we have

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, A\mathbf{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, A\mathbf{z} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then find the value of the determinant of the matrix $A$.

## Solution 1.

Let $B$ be the $3 \times 3$ matrix whose columns are the vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$, that is,

$$B = [\mathbf{x}\mathbf{y}\mathbf{z}].$$

Then we have

$$AB = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Then we have

$$\det(A)\det(B) = \det(AB) = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 0.$$

(If two rows are equal, then the determinant is zero. Or you may compute the determinant by the second column cofactor expansion.)

Note that the column vectors of $B$ are linearly independent, and hence $B$ is nonsingular matrix. Thus the $\det(B) \neq 0$.
Therefore the determinant of $A$ must be zero.

## Solution 2.

Since

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

we have

$$A\mathbf{x} + A\mathbf{y} = A\mathbf{z}.$$

It follows that we have

$$A(\mathbf{x} + \mathbf{y} - \mathbf{z}) = \mathbf{0}.$$

Since the vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are linearly independent, the linear combination $\mathbf{x} + \mathbf{y} - \mathbf{z} \neq \mathbf{0}$.

Hence the matrix $A$ is singular, and the determinant of $A$ is zero.

(Recall that a matrix $A$ is singular if and only if there exist nonzero vector $\mathbf{v}$ such that $A\mathbf{u} = \mathbf{0}$.)

---

13 Determine the values of $x$ so that the matrix

$$A = \begin{bmatrix} 1 & 1 & x \\ 1 & x & x \\ x & x & x \end{bmatrix}$$

is invertible. For those values of $x$, find the inverse matrix $A^{-1}$.

### Solution.

We use the fact that a matrix is invertible if and only if its determinant is nonzero.

So we compute the determinant of the matrix $A$.

We have

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 1 & x \\ 1 & x & x \\ x & x & x \end{vmatrix} \\ &= (1)\begin{vmatrix} x & x \\ x & x \end{vmatrix} - (1)\begin{vmatrix} 1 & x \\ x & x \end{vmatrix} + x\begin{vmatrix} 1 & x \\ x & x \end{vmatrix} && \text{by the first row cofactor expansion.} \\ &= (x^2 - x^2) - (x - x^2) + x(x - x^2) \\ &= (x-1)(x - x^2) \\ &= x(x-1)^2. \end{aligned}$$

Thus, the determinant $\det(A)$ is zero if and only if $x = 0, 1$.

Hence the matrix $A$ is invertible if and only if $x \neq 0, 1$.

Next, we suppose that $x \neq 0, 1$ and find the inverse matrix of $A$.

We reduce the augmented matrix $[A \mid I]$ as follows.

We have

$$[A \mid I] = \begin{bmatrix} 1 & 1 & x & 1 & 0 & 0 \\ 1 & x & x & 0 & 1 & 0 \\ x & x & x & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_3 - xR_1]{R_2 - R_1} \begin{bmatrix} 1 & 1 & x & 1 & 0 & 0 \\ 0 & x-1 & 0 & -1 & 1 & 0 \\ 0 & 0 & x-x^2 & -x & 0 & 1 \end{bmatrix} \xrightarrow[\frac{1}{x-x^2}R_3]{\frac{1}{x-1}R_2} \begin{bmatrix} 1 & 1 & x & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{-1}{x-1} & \frac{1}{x-1} & 0 \\ 0 & 0 & 1 & \frac{-1}{1-x} & 0 & \frac{1}{x-x^2} \end{bmatrix}$$

$$\xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 0 & x & \frac{x}{x-1} & \frac{-1}{x-1} & 0 \\ 0 & 1 & 0 & \frac{-1}{x-1} & \frac{1}{x-1} & 0 \\ 0 & 0 & 1 & \frac{-1}{1-x} & 0 & \frac{1}{x-x^2} \end{bmatrix} \xrightarrow{R_1 - xR_3} \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{-1}{x-1} & \frac{-x}{x-x^2} \\ 0 & 1 & 0 & \frac{-1}{x-1} & \frac{1}{x-1} & 0 \\ 0 & 0 & 1 & \frac{-1}{1-x} & 0 & \frac{1}{x-x^2} \end{bmatrix}.$$

Now that we reduced the left $3 \times 3$ matrix into the identity matrix, the right $3 \times 3$ matrix is the inverse matrix of $A$.

(Note that when we applied elementary row operations, we divided by $x - 1$ and $x - x^2$, and this is where we needed to assume $x \neq 0, 1$.)

We have

$$A^{-1} = \begin{bmatrix} 0 & \frac{-1}{x-1} & \frac{-x}{x-x^2} \\ \frac{-1}{x-1} & \frac{1}{x-1} & 0 \\ \frac{-1}{1-x} & 0 & \frac{1}{x-x^2} \end{bmatrix} = \frac{1}{x(1-x)} \begin{bmatrix} 0 & x & -x \\ x & -x & 0 \\ -x & 0 & 1 \end{bmatrix}.$$

---

14 Prove that if $n \times n$ matrices $A$ and $B$ are nonsingular, then the product $AB$ is also a nonsingular matrix.
(*The Ohio State University*)

### Definition (Nonsingular Matrix)

An $n \times n$ matrix is called nonsingular if the only solution $\mathbf{x} \in \mathbb{R}^n$ of the equation $A\mathbf{x} = \mathbf{x}$ is $\mathbf{x} = \mathbf{0}$.

### Proof.

We give two proofs. The first one uses a property of the determinants of matrices, and the second one uses the definition of nonsingular matrices.

### Proof 1. (Using Determinant)

Recall that a matrix is nonsingular if and only if its determinant is not zero.

Since $A$ and $B$ are nonsingular, we know that

$$\det(A) \neq 0 \text{ and } \det(B) \neq 0.$$

Then we have using the multiplicative property of the determinant

$$\det(AB) = \det(A)\det(B) \neq 0.$$

Since the determinant of the product $AB$ is not zero, we conclude that $AB$ is a nonsingular matrix.

### Proof 2. (Using Definition of Nonsingular Matrices)

Suppose that $A, B$ are nonsingular matrices.
This means that if $A\mathbf{x} = \mathbf{0}$ for some the vector $\mathbf{x} \in \mathbb{R}^n$, then we must have $\mathbf{x} = \mathbf{0}$.
Same for $B$.

Suppose that we have $(AB)\mathbf{x} = \mathbf{0}$ for some vector $\mathbf{x} \in \mathbb{R}^n$.
Let $\mathbf{v} = B\mathbf{x} \in \mathbb{R}^n$. Then we have

$$A\mathbf{v} = AB\mathbf{x} = \mathbf{0}.$$

So since $A$ is a nonsingular matrix, we have $\mathbf{v} = \mathbf{0}$, namely, $B\mathbf{x} = \mathbf{0}$.

Since $B$ is nonsingular, this further implies that $\mathbf{x} = \mathbf{0}$.

In summary, whenever $(AB)\mathbf{x} = \mathbf{0}$, we have $\mathbf{x} = \mathbf{0}$.
Therefore, the matrix $AB$ is nonsingular.

⑮ Let $n$ be an odd positive integer. Determine whether there exists an $n \times n$ real matrix $A$ such that $A^2 + I = O$, where $I$ is the $n \times n$ identity matrix and $O$ is the $n \times n$ zero matrix. If such a matrix $A$ exists, find an example. If not, prove that there is no such $A$. How about when $n$ is an even positive number?

### Hint.

The key technique to solve this problem very easily is determinant.
Recall the following properties of the determinant.
Let $A, B$ be $n \times n$ matrices and $c$ be a scalar Then we have

⑯ $\det(AB) = \det(A)\det(B)$

⑰ $\det(cA) = c^n \det(A)$.

## Solution.

### When $n$ is odd.

When $n$ is odd, we prove that there is no $n \times n$ real matrix $A$ such that $A^2 + I = O$.
Seeking a contradiction, assume that we have $A$ such that $A^2 + I = O$.
Since we have $A^2 = -I$, we have

$$\det(A^2) = \det(-I).$$

Using the properties of determinant, we obtain

$$\det(A)^2 = (-1)^n \det(I) = -1$$

because $n$ is odd and $\det(I) = 1$.

Since $A$ is a real matrix, the determinant of $A$ is also real.
Thus, $\det(A)^2 = -1$ is impossible. Hence there is no such $A$.

### When $n$ is even.

On the other hand, if $n$ is even there is $A$ such that $A^2 + I = O$.
For example, consider

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Then a direct computation shows that $A^2 = -I$, hence $A^2 + I = O$.

## Comment.

Recall that the imaginary number $i$ is the number whose square is $-1$.
Similarly, we found above that the square of the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is $-I$.

But when $n$ is odd, there is no such matrix $A$ as we showed.

---

18 Prove that the determinant of an $n \times n$ skew-symmetric matrix is zero if $n$ is odd.

## Definition (Skew-Symmetric)

A matrix $A$ is called **skew-symmetric** if $A^{\mathrm{T}} = -A$.

Here $A^{\mathrm{T}}$ is the transpose of $A$.

## Proof.

### Properties of Determinants

We will use the following two properties of determinants of matrices.
For any $n \times n$ matrix $A$ and a scalar $c$, we have

⑲ $\det(A) = \det(A^{\mathrm{T}})$,

⑳ $\det(cA) = c^n \det(A)$.

## Main Part of the Proof

Suppose that $n$ is an odd integer and let $A$ be an $n \times n$ skew-symmetric matrix.
Thus, we have

$$A^{\mathrm{T}} = -A$$

by definition of skew-symmetric.

Then we have

$$\begin{aligned}
\det(A) &= \det(A^{\mathrm{T}}) && \text{by property 1} \\
&= \det(-A) && \text{since } A \text{ is skew-symmetric} \\
&= (-1)^n \det(A) && \text{by property 2} \\
&= -\det(A) && \text{since } n \text{ is odd.}
\end{aligned}$$

Therefore, it yields that $2\det(A) = 0$, and hence $\det(A) = 0$.

## Comment.

The result implies that every odd degree skew-symmetric matrix is not invertible, or equivalently singular.

Also, this means that each odd degree skew-symmetric matrix has the eigenvalue $0$.

## Related Question.

The eigenvalues of a real skew-symmetric matrices are of the special form as in the next problem.

**Problem. (a)** Prove that each eigenvalue of the real skew-symmetric matrix $A$ is either $0$ or a purely imaginary number.
**(b)** Prove that the rank of $A$ is even.

# Computations of Determinants

## Problems

1 Find the determinant of the matix

$$A = \begin{bmatrix} 100 & 101 & 102 \\ 101 & 102 & 103 \\ 102 & 103 & 104 \end{bmatrix}.$$

## Solution.

Note that the determinant does not change if the $i$-th row is added by a scalar multiple of the $j$-th row if $i \neq j$.
We use this fact about the determinant and compute $\det(A)$ as follows.

$$\begin{aligned} \det(A) &= \begin{vmatrix} 100 & 101 & 102 \\ 101 & 102 & 103 \\ 102 & 103 & 104 \end{vmatrix} \\ &= \begin{vmatrix} 100 & 101 & 102 \\ 101 & 102 & 103 \\ 1 & 1 & 1 \end{vmatrix} \quad (\text{by } R_3 - R_2) \\ &= \begin{vmatrix} 100 & 101 & 102 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad (\text{by } R_2 - R_1) \\ &= \begin{vmatrix} 100 & 101 & 102 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{vmatrix} \quad (\text{by } R_3 - R_1) \\ &= 0 \quad (\text{by the third row cofactor expansion.}) \end{aligned}$$

Therefore the determinant $\det(A)$ is zero.

# Introduction to Eigenvalues and Eigenvectors

**Definition**

Let $A$ be an $n \times n$ matrix.

1. A scalar $\lambda$ is called an **eigenvalue** of $A$ if the equation $A\mathbf{x} = \lambda\mathbf{x}$ has a nonzero solution $\mathbf{x}$. Such a nonzero solution $\mathbf{x}$ is called an **eigenvector** corresponding to the eigenvalue $\lambda$.
2. The **characteristic polynomial** of $A$ is the degree $n$ polynomial $p(t) = \det(A - tI)$.
3. If $p(t) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_k)^{n_k}$ is a factorization of the characteristic polynomial of $A$, where $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues of $A$, then the **algebraic multiplicity** of the eigenvalue $\lambda_i$ is $n_i$.

**Summary**

Let $A$ be an $n \times n$ matrix. Let $p(t)$ be the characteristic polynomial of $A$.

1. The degree of $p(t)$ is $n$.
2. $\lambda$ is an eigenvalue of $A$ if and only if $p(\lambda) = \det(A - \lambda I) = 0$.
3. $A$ has at least one eigenvalue and has at most $n$ distinct eigenvalues.
4. $A$ has at most $n$ distinct eigenvalues.
5. The eigenvalues of a matrix $A$ are roots of the characteristic polynomial of $A$.
6. The eigenvalues of a triangular matrix are diagonal entries.

## Problems

1. **(a)** True or False. If each entry of an $n \times n$ matrix $A$ is a real number, then the eigenvalues of $A$ are all real numbers.
**(b)** Find the eigenvalues of the matrix

$$B = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}.$$

(*The Ohio State University*)

**Hint.**

Consider a $2 \times 2$ matrix.

Then the eigenvalues are solutions of a quadratic polynomial.

Does a quadratic polynomial always have real solutions?

**Solution.**

(a) True or False. If each entry of an $n \times n$ matrix $A$ is a real number, then the eigenvalues of $A$ are all real numbers.

False. In general, a real matrix can have a complex number eigenvalue. In fact, the part (b) gives an example of such a matrix.

(b) Find the eigenvalues of the matrix

The characteristic polynomial for $B$ is

$$\det(B - tI) = \begin{bmatrix} -2-t & -1 \\ 5 & 2-t \end{bmatrix} = t^2 + 1.$$

The eigenvalues are the solutions of the characteristic polynomial. Thus solving $t^2 + 1 = 0$, we obtain eigenvalues $\pm i$, where $i = \sqrt{-1}$.

Thus the eigenvalue of a real matrix $B$ is pure imaginary numbers $\pm i$.

---

② Find all the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 3 & -2 \\ 6 & -4 \end{bmatrix}$.

### (a) Solution.

To determine eigenvalues of $A$, we compute the determinant of $A - \lambda I$.

We have

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & -2 \\ 6 & -4-\lambda \end{vmatrix} \\ &= (3-\lambda)(-4-\lambda) + 12 \\ &= \lambda^2 + \lambda = \lambda(\lambda + 1). \end{aligned}$$

The eigenvalues are solutions of $\det(A - \lambda I) = 0$, hence eigenvalues of $A$ are $0, -1$.

---

Next, we find the eigenvector corresponding to the eigenvalue $\lambda = 0$.

Eigenvectors $\mathbf{x}$ are nonzero solutions of $(A - 0I)\mathbf{x} = \mathbf{0}$.

Thus, we solve $A\mathbf{x} = \mathbf{0}$. The augmented matrix of the system is

$$\left[\begin{array}{cc|c} 3 & -2 & 0 \\ 6 & -4 & 0 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{cc|c} 3 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{cc|c} 1 & -2/3 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Thus, if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a solution, then $x_1 = \frac{2}{3}x_2$, hence

$$\mathbf{x} = x_2 \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to $\lambda = 0$ for any nonzero scalar $x_2$.

Finally, we find the eigenvectors corresponding to the eigenvalue $\lambda = -1$.

In this case we need to solve $(A - (-1)I)\mathbf{x} = \mathbf{0}$.

The augmented matrix is

$$\left[\begin{array}{cc|c} 4 & -2 & 0 \\ 6 & -3 & 0 \end{array}\right] \xrightarrow[\frac{1}{3}R_2]{\frac{1}{2}R_1} \left[\begin{array}{cc|c} 2 & -1 & 0 \\ 2 & -1 & 0 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{2}R_1} \left[\begin{array}{cc|c} 1 & -1/2 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

It follows that the eigenvectors associated to $\lambda = -1$ are

$$\mathbf{x} = x_2 \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

for any nonzero scalar $x_2$.

③ Let

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Find the eigenvalues of the matrix $A$. Also give the algebraic multiplicity of each eigenvalue.

### Solution (a). Eigenvalues of $A$ and algebraic multiplies

Eigenvalues and their algebraic multiplicities are determined by the characteristic polynomial $p(t)$ of $A$.

By definition, the characteristic polynomial of $A$ is $p(t) = \det(A - tI)$.

We have

$$\begin{aligned} p(t) &= \det(A - tI) \\ &= \begin{vmatrix} -t & 0 & 0 & 0 \\ 1 & 1-t & 1 & 1 \\ 0 & 0 & -t & 0 \\ 1 & 1 & 1 & 1-t \end{vmatrix} \\ &= -t \begin{vmatrix} 1-t & 1 & 1 \\ 0 & -t & 0 \\ 1 & 1 & 1-t \end{vmatrix} && \text{by the first row cofactor expansion} \\ &= -t\left(-t \begin{vmatrix} 1-t & 1 \\ 1 & 1-t \end{vmatrix}\right) && \text{by the second row cofactor expansion} \\ &= t^2\left((1-t)^2 - 1\right) \\ &= t^2(t^2 - 2t) \\ &= t^3(t-2). \end{aligned}$$

Thus the characteristic polynomial is

$$p(t) = t^3(t-2).$$

From this, the eigenvalues of $A$ are 0 and 2 with algebraic multiplicities 3 and 1, respectively.

## Solution (b). Geometric multiplicity

We give two solutions for part (b).

### First Solution (b). (Finding the rank first)

Recall that the geometric multiplicity of $\lambda$ is the dimension of the eigenspace $E_\lambda = \mathcal{N}(A-\lambda I)$. That is, the geometric multiplicity of $\lambda$ is the nullity of the matrix $A - \lambda I$.

Let us now consider the case $\lambda = 0$.

We first find the rank of $A - 0I = A$ as follows.

$$A - 0I = A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_4 - R_2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is in reduced row echelon form.

Hence the rank of $A$ is 1.

The rank-nullity theorem says that

$$\text{rank of } A \ + \ \text{nullity of } A = 4.$$

Thus, the nullity of $A = A - 0I$ is 3, and hence the geometric multiplicity of $\lambda = 0$ is 3.

### Second Solution (b). (Finding a basis of the eigenspace)

In this solution, we find a basis of the eigenspace $E_0$.

By definition $E_0 = \mathcal{N}(A - 0I) = \mathcal{N}(A)$.

Thus, the eigenspace $E_0$ is the null space of the matrix $A$.

We solve the equation $A\mathbf{x} = \mathbf{0}$ as follows.

The augmented matrix of this equation is

$$[A \mid \mathbf{0}] = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_4 - R_2} \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Hence the solution satisfies

$$x_1 = -x_2 - x_3 - x_4$$

and the general solution is

$$\mathbf{x} = \begin{bmatrix} -x_2 - x_3 - x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, the eigenspace is

$$\begin{aligned} E_0 &= \mathcal{N}(A) \\ &= \left\{ \mathbf{x} \in \mathbb{C}^4 \;\middle|\; \mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \text{ for any } x_2, x_3, x_4 \in \mathbb{C} \right\} \\ &= \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Thus the set

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a spanning set of $E_0$, and it is straightforward to check that the set is linearly independent.
Hence this set is a basis of $E_0$, and the dimension of $E_0$ is 3.
The geometric multiplicity of $\lambda = 0$ is the dimension of $E_0$ by definition.
Thus, the geometric multiplicity of $\lambda$ is 3.

---

4 Let

$$A = \begin{bmatrix} a & -1 \\ 1 & 4 \end{bmatrix}$$

be a $2 \times 2$ matrix, where $a$ is some real number. Suppose that the matrix $A$ has an eigenvalue $3$.

**(a)** Determine the value of $a$.
**(b)** Does the matrix $A$ have eigenvalues other than $3$?

## Solution.

### (a) Determine the value of $a$

Since $3$ is an eigenvalue of the matrix $A$, we have

$$0=\det(A-3I),$$

where $I$ is the $2\times 2$ identity matrix.

Thus we have

$$\begin{aligned}0&=\det(A-3I)\\&=\begin{vmatrix}a-3 & -1\\ 1 & 4-3\end{vmatrix}\\&=\begin{vmatrix}a-3 & -1\\ 1 & 1\end{vmatrix}\\&=(a-3)(1)-(-1)(1)=a-2.\end{aligned}$$

Thus the value of $a$ must be $2$.

(b) Does the matrix $A$ have eigenvalues other than $3$?

Let us determine all the eigenvalue of the matrix

$$A=\begin{bmatrix}2 & -1\\ 1 & 4\end{bmatrix}.$$

We compute the characteristic polynomial $p(t)=\det(A-tI)$ of $A$.

We have

$$\begin{aligned}p(t)&=\det(A-tI)\\&=\begin{vmatrix}2-t & -1\\ 1 & 4-t\end{vmatrix}\\&=(2-t)(4-t)-(-1)(1)\\&=t^2-6t+9\\&=(t-3)^2.\end{aligned}$$

Since the eigenvalues are roots of the characteristic polynomial, solving $(t-3)^2=0$ we see that $t=3$ is the only eigenvalue of $A$ (with algebraic multiplicity $2$).

Hence the matrix $A$ does not have eigenvalues other than $3$.

---

⑤ Determine all eigenvalues and their algebraic multiplicities of the matrix

$$A=\begin{bmatrix}1 & a & 1\\ a & 1 & a\\ 1 & a & 1\end{bmatrix},$$

where $a$ is a real number.

Proof.

(a) $\lambda+c$ is an eigenvalue of $A+cI$.

Let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$. Then we have

$$A\mathbf{x}=\lambda\mathbf{x}.$$

It follows that we have

$$\begin{aligned}(A+cI)\mathbf{x} &= A\mathbf{x}+c\mathbf{x}\\ &= \lambda\mathbf{x}+c\mathbf{x}\\ &= (\lambda+c)\mathbf{x}.\end{aligned}$$

Thus, we obtain

$$(A+cI)\mathbf{x}=(\lambda+c)\mathbf{x},$$

where $\mathbf{x}$ is a nonzero vector.
Hence $\lambda+c$ is an eigenvalue of the matrix $A+cI$, and $\mathbf{x}$ is an eigenvector corresponding to $\lambda-c$.

In summary, if $\lambda$ is an eigenvalue of $A$ and $\mathbf{x}$ is an associated eigenvector, then $\lambda+c$ is an eigenvalue of $A+cI$ and $\mathbf{x}$ is an associated eigenvector corresponding to $\lambda+c$.

### (b) Algebraic multiplicities are the same

Let $p(t)=\det(A-tI)$ be the characteristic polynomial of $A$.
Let $q(t)$ be the characteristic polynomial of the matrix $A+cI$.
Then we have

$$\begin{aligned}q(t) &= \det\left((A+cI)-t\right)=\det\left(A-(t-c)\right)\\ &= p(t-c).\end{aligned}$$

Let $\lambda_1,\dots,\lambda_k$ be distinct eigenvalues of $A$ with algebraic multiplicities $n_1,\dots,n_k$, respectively.
Then we have

$$p(t)=\pm\prod_{i=1}^{k}(t-\lambda_i)^{n_i}.$$

It follows that we have

$$\begin{aligned}q(t) &= p(t-c)\\ &= \pm\prod_{i=1}^{k}(t-c-\lambda_i)^{n_i}\\ &= \pm\prod_{i=1}^{k}\left(t-(\lambda_i+c)\right)^{n_i}.\end{aligned}$$

From the last equation, we read that the eigenvalues of the matrix $A+cI$ are $\lambda_i+c$ with algebraic multiplicity $n_i$ for $i=1,\dots,k$.
Thus, geometric multiplicities of $\lambda$ and $\lambda-c$ are the same.

### (c) How about geometric multiplicities?

From part (a), we know that eigenvectors of $\lambda$ are eigenvectors of $\lambda-c$.
Reversing the argument, the eigenvectors of $\lambda+c$ are eigenvectors of $\lambda$.

Thus, the eigenvectors correspond one to one, and the eigenspace of $\lambda$ is the same as the eigenspace of $\lambda+c$.

Hence their geometric multiplicities are the same.

## Applications

As applications of this problem, consider the following problems.

*Problem*. Find all the eigenvalues and eigenvectors of the matrix

$$A=\begin{bmatrix} 3 & 9 & 9 & 9 \\ 9 & 3 & 9 & 9 \\ 9 & 9 & 3 & 9 \\ 9 & 9 & 9 & 3 \end{bmatrix}.$$

This is a problem of a Linear Algebra final exam at Harvard University.
For a solution of this problem, see the post "Eigenvalues and eigenvectors of matrix whose diagonal entries are 3 and 9 elsewhere".

*Problem*. Find the determinant of the following matrix

$$A=\begin{bmatrix} 6 & 2 & 2 & 2 & 2 \\ 2 & 6 & 2 & 2 & 2 \\ 2 & 2 & 6 & 2 & 2 \\ 2 & 2 & 2 & 6 & 2 \\ 2 & 2 & 2 & 2 & 6 \end{bmatrix}.$$

This is also a problem of a Linear Algebra final exam at Harvard University.
See the post "Determinant of matrix whose diagonal entries are 6 and 2 elsewhere" for a solution.

---

⑥ Suppose that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of a matrix $A$ corresponding to the eigenvalue $3$ and that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector of $A$ corresponding to the eigenvalue $-2$. Compute $A^2\begin{bmatrix} 4 \\ 3 \end{bmatrix}$.

(*Stanford University*)

## Solution.

By the given conditions, we have

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix}=3\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } A\begin{bmatrix} 2 \\ 1 \end{bmatrix}=-2\begin{bmatrix} 2 \\ 1 \end{bmatrix}. \tag{*}$$

We express the vector $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as a linear combination of vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Let

$$a\begin{bmatrix} 1 \\ 1 \end{bmatrix}+b\begin{bmatrix} 2 \\ 1 \end{bmatrix}=\begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

and we want to determine scalars $a$ and $b$. The augmented matrix of the system is reduced as follows:

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 1 & 1 & 3 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right].$$

Thus the solution is

$$a = 2, b = 1$$

and hence we obtain the linear combination

$$2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

---

Before computing $A^2 \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, we first calculate $A \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as follows.

$$\begin{aligned} A\begin{bmatrix} 4 \\ 3 \end{bmatrix} &= A\left(2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) \\ &= 2A\begin{bmatrix} 1 \\ 1 \end{bmatrix} + A\begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= 2 \cdot 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-2)\begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ by } (*) \\ &= 6\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 1 \end{bmatrix}. \end{aligned}$$

---

Now, using the result of the above computation, we compute

$$\begin{aligned} A^2\begin{bmatrix} 4 \\ 3 \end{bmatrix} &= A\left(\begin{bmatrix} 4 \\ 3 \end{bmatrix}\right) \\ &= A\left(6\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) \\ &= 6A\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2A\begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= 6 \cdot 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2(-2)\begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ by } (*) \\ &= \begin{bmatrix} 26 \\ 22 \end{bmatrix}. \end{aligned}$$

Therefore we obtained

$$A^2\begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 26 \\ 22 \end{bmatrix}.$$

7 Suppose that $A$ is an $n \times n$ matrix with eigenvalue $\lambda$ and corresponding eigenvector $\mathbf{v}$.

**(a)** If $A$ is invertible, is $\mathbf{v}$ an eigenvector of $A^{-1}$? If so, what is the corresponding eigenvalue? If not, explain why not.

**(b)** Is $3\mathbf{v}$ an eigenvector of $A$? If so, what is the corresponding eigenvalue? If not, explain why not.

(*Stanford University*)

Hint.

Use the defining relation $A\mathbf{v} = \lambda\mathbf{v}$.

## Solution.

### (a) If $A$ is invertible, is $\mathbf{v}$ an eigenvector of $A^{-1}$?

The answer is yes. First note that the eigenvalue $\lambda$ is not zero since $A$ is invertible.

By definition, we have $A\mathbf{v} = \lambda\mathbf{v}$. Multiplying it by $A^{-1}$ from the left, we have

$$\mathbf{v} = \lambda A^{-1}\mathbf{v}.$$

As noted above, $\lambda$ is not zero, so we divide this equality by $\lambda$ and obtain

$$A^{-1}\mathbf{v} = \frac{1}{\lambda}\mathbf{v}.$$

Since $\mathbf{v}$ is not a zero vector, this implies that $\mathbf{v}$ is an eigenvector corresponding to the eigenvalue $1/\lambda$ of $A$.

### (b) Is $3\mathbf{v}$ an eigenvector of $A$?

The answer is yes. We calculate

$$A(3\mathbf{v}) = 3A\mathbf{v} = 3\lambda\mathbf{v} = \lambda(3\mathbf{v}).$$

Thus we have $A(3\mathbf{v}) = \lambda(3\mathbf{v})$.

Since $3\mathbf{v} \neq 0$, this implies that $3\mathbf{v}$ is an eigenvector corresponding to the eigenvalue $\lambda$.

---

8 Let $A$ be a $2 \times 2$ real symmetric matrix. Prove that all the eigenvalues of $A$ are real numbers by considering the characteristic polynomial of $A$.

## Proof.

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Then as $A$ is a symmetric matrix, we have $A^{\mathrm{T}} = A$.

This implies that

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Hence we have $b = c$ by comparing entries.

---

Now, we find the characteristic polynomial $p(t)$ of $A$.

We have

$$p(t) = \det(A - tI) = \begin{vmatrix} a-t & b \\ b & d-t \end{vmatrix}$$

$$\begin{aligned} &= (a-t)(d-t) - b^2 \\ &= t^2 - (a+d)t + ad - b^2. \end{aligned}$$

Note that the eigenvalues of $A$ are roots of the characteristic polynomial $p(t)$. Hence, it suffices to show that the roots of $p(t)$ are real numbers.

The quadratic polynomial has only real roots if and only if its discriminant is non-negative.

The discriminant of $p(t)$ is given by

$$\begin{aligned} (a+d)^2 - 4(ad - b^2) &= a^2 + 2ad + d^2 - 4ad + 4b^2 \\ &= a^2 - 2ad + d^2 + 4b^2 \\ &= (a-d)^2 + 4b^2. \end{aligned}$$

Observe that the last expression is the sum of two squares of real numbers. Hence the discriminant of $p(t)$ is nonnegative.

We conclude that every 2 × 2 symmetric matrix has only real eigenvalues.

**Remark**

We also could find the eigenvalues directly. By the quadratic formula, the eigenvalues of $A$ are

$$\frac{a + d \pm \sqrt{(a+d)^2 - 4(ad - b^2)}}{2} = \frac{a + d \pm \sqrt{(a-d)^2 + 4b^2}}{2}$$

and as the number inside the square root (discriminant) is positive, we conclude that the eigenvalues are real.

---

**9** Let

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

be a $2 \times 2$ matrix, where $a, b$ are real numbers. Suppose that $b \neq 0$. Prove that the matrix $A$ does not have real eigenvalues.

**Proof.**

Let $\lambda$ be an arbitrary eigenvalue of $A$.

Then the matrix $A - \lambda I$ is singular, where $I$ is the $2 \times 2$ identity matrix.

This is equivalent to having $\det(A - \lambda I) = 0$.

We compute the determinant as follows.

We have

$$\begin{aligned}\det(A-\lambda I) &= \begin{vmatrix} a-\lambda & b \\ -b & a-\lambda \end{vmatrix} \\ &= (a-\lambda)^2 - b(-b) \\ &= a^2 - 2a\lambda + \lambda^2 + b^2 \\ &= \lambda^2 - 2a\lambda + a^2 + b^2.\end{aligned}$$

We solve the equation $\lambda^2 - 2a\lambda + a^2 + b^2 = 0$ by the quadratic formula and obtain

$$\begin{aligned}\lambda &= \frac{2a \pm \sqrt{4a^2 - 4(a^2+b^2)}}{2} = \frac{2a \pm \sqrt{-4b^2}}{2} \\ &= a \pm |b|i.\end{aligned}$$

Since $b \neq 0$ by assumption, the eigenvalue $\lambda = a \pm |b|i$ is not a real number.

As $\lambda$ is an arbitrary eigenvalue of $A$, we conclude that all eigenvalues of $A$ are not real numbers.

---

**10** Let

$$A = \begin{bmatrix} 3 & -12 & 4 \\ -1 & 0 & -2 \\ -1 & 5 & -1 \end{bmatrix}.$$

Then find all eigenvalues of $A^5$. If $A$ is invertible, then find all the eigenvalues of $A^{-1}$.

### Proof.

We first determine all the eigenvalues of the matrix $A$.

The characteristic polynomial $p(t)$ of $A$ is given by

$$\begin{aligned}p(t) &= \det(A - tI) \\ &= \begin{vmatrix} 3-t & -12 & 4 \\ -1 & -t & -2 \\ -1 & 5 & -1-t \end{vmatrix}.\end{aligned}$$

Using the first row cofactor expansion, we compute

$$\begin{aligned}p(t) &= (3-t)\begin{vmatrix} -t & -2 \\ 5 & -1-t \end{vmatrix} - (-12)\begin{vmatrix} -1 & -2 \\ -1 & -1-t \end{vmatrix} + 4\begin{vmatrix} -1 & -t \\ -1 & 5 \end{vmatrix} \\ &= (3-t)(t^2+t+10) + 12(t-1) + 4(-5-t) \\ &= -t^3 + 2t^2 + 8t - 2.\end{aligned}$$

Therefore the characteristic polynomial of $A$ is

$$p(t) = -t^3 + 2t^2 + 8t - 2$$

and it can be factored as

$$p(t) = -(t-2)(t-1)(t+1).$$

The roots of the characteristic polynomials are all the eigenvalues of $A$.

Thus, $2, \pm 1$ are the eigenvalues of $A$.

---

To find the eigenvalues of $A^5$, recall that if $\lambda$ is an eigenvalue of $A$, then $\lambda^5$ is an eigenvalue of $A^5$. It follows from this fact that $2^5, (-1)^5, 1^5$ are eigenvalues of $A^5$.

Since $A^5$ is a $3 \times 3$ matrix, its characteristic polynomial has degree $3$, hence there are at most $3$ distinct eigenvalues of $A^5$.

Because we have found three eigenvalues, $32, -1, 1$, of $A^5$, these are all the eigenvalues of $A^5$.

---

Recall that a matrix is singular if and only if $\lambda = 0$ is an eigenvalue of the matrix.

Since $0$ is not an eigenvalue of $A$, it follows that $A$ is nonsingular, and hence invertible. If $\lambda$ is an eigenvalue of $A$, then $\frac{1}{\lambda}$ is an eigenvalue of the inverse $A^{-1}$.

So $\frac{1}{\lambda}, \lambda = 2, \pm 1$ are eigenvalues of $A^{-1}$.

As above, the matrix $A^{-1}$ is $3 \times 3$, hence it has at most three distinct eigenvalues. We have found $1/2, \pm 1$ are eigenvalues of $A^{-1}$, hence these are all the eigenvalues of $A^{-1}$.

---

In summary, all the eigenvalues of $A^5$ are $\pm 1, 32$. The matrix $A$ is invertible and all the eigenvalues of $A^{-1}$ are $\pm 1, 1/2$.

### Comment.

Do not try to compute $A^5$ and $A^{-1}$ and then find their eigenvalues.

It will be tedious for hand computation.

---

11 Let $\lambda$ be an eigenvalue of $n \times n$ matrices $A$ and $B$ corresponding to the same eigenvector $\mathbf{x}$.

**(a)** Show that $2\lambda$ is an eigenvalue of $A + B$ corresponding to $\mathbf{x}$.

**(b)** Show that $\lambda^2$ is an eigenvalue of $AB$ corresponding to $\mathbf{x}$.

(*The Ohio State University*)

### Proof.

#### (a) Show that $2\lambda$ is an eigenvalue of $A + B$ corresponding to $\mathbf{x}$.

Since $\lambda$ is an eigenvalue of $A$ and $B$, and $\mathbf{x}$ is a corresponding eigenvector, we have

$$A\mathbf{x} = \lambda\mathbf{x} \text{ and } B\mathbf{x} = \lambda\mathbf{x}.$$

Then we compute

$$\begin{aligned}(A + B)\mathbf{x} &= A\mathbf{x} + B\mathbf{x} \\ &= \lambda\mathbf{x} + \lambda\mathbf{x} \qquad \text{by } (*) \\ &= 2\lambda\mathbf{x}.\end{aligned}$$

Since $\mathbf{x}$ is an eigenvector, it is a nonzero vector by definition.

Hence from the equality

$$(A+B)\mathbf{x} = 2\lambda\mathbf{x},$$

we see that $2\lambda$ is an eigenvalue of the matrix $A+B$ and $\mathbf{x}$ is an associated eigenvector.

(b) Show that $\lambda^2$ is an eigenvalue of $AB$ corresponding to $\mathbf{x}$.

We have

$$\begin{aligned}(AB)\mathbf{x} &= A(B\mathbf{x}) \\ &= A(\lambda\mathbf{x}) && \text{by } (*) \\ &= \lambda(A\mathbf{x}) \\ &= \lambda(\lambda\mathbf{x}) && \text{by } (*) \\ &= \lambda^2\mathbf{x}.\end{aligned}$$

Since $\mathbf{x}$ is a nonzero vector as it is an eigenvector, it follows from the equality

$$(AB)\mathbf{x} = \lambda^2\mathbf{x}$$

that $\lambda^2$ is an eigenvalue of the matrix $AB$ and $\mathbf{x}$ is a corresponding eigenvector.

---

12 Find all eigenvalues of the following $n \times n$ matrix.

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

## Steps.

13 Use the 1st row cofactor expansion

14 For each smaller determinant matrix, use cofactor expansion inductively.

## Solution.

We calculate the characteristic polynomial $\det(A-\lambda I)$ of $A$.

We use the cofactor expansion corresponding to the first row.

$$\det(A-\lambda I)=\begin{vmatrix} -\lambda & 0 & \cdots & 0 & 1 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{vmatrix}$$

$$=-\lambda\begin{vmatrix} -\lambda & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{vmatrix}+(-1)^{n+1}\begin{vmatrix} 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

Now we calculate the two determinants in the second equality separately.

The easiest way is to note that the determinant of a triangular matrix is the product of its diagonal entries.
Thus the first determinant is $(-\lambda)^{n-1}$ and the second determinant is $1$.

(If you don't know this fact, then use the first row cofactor expansion inductively to compute the first determinant. For the second one, use the first column cofactor expansion inductively.)

Thus we obtain $\det(A-\lambda I)=(-1)^n\lambda^n+(-1)^{n+1}=(-1)^n(\lambda^n-1)$.
Therefore eigenvalues are $n$-th roots of unity $e^{2\pi i/n}$ for $i=0,1,\dots,n-1$.

## Comment.

The original determinant is not in a good shape for induction but once we apply the 1st row cofactor expansion the smaller determinants obtained are better suited for induction.
When I say in the proof "inductively", I meant that you need to use mathematical induction to prove the claim (more) rigorously.

15

Find all the eigenvalues of the matrix $A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

(*The Ohio State University*)

## Solution.

We compute the characteristic polynomial $p(t)$ of the matrix $A$ as follows.
We have

$$p(t) = \det(A - tI)$$

$$= \begin{vmatrix} -t & 1 & 0 & 0 \\ 0 & -t & 1 & 0 \\ 0 & 0 & -t & 1 \\ 1 & 0 & 0 & -t \end{vmatrix}$$

$$= -t \begin{vmatrix} -t & 1 & 0 \\ 0 & -t & 1 \\ 0 & 0 & -t \end{vmatrix} - \begin{vmatrix} 0 & 1 & 0 \\ 0 & -t & 1 \\ 1 & 0 & -t \end{vmatrix}$$

by the first row cofactor expansion.

The left determinant of the $3 \times 3$ matrix in (*) is $(-t)^3$ since it is a diagonal matrix.

We apply the first column cofactor expansion to the right determinant in (*) and obtain

$$\begin{vmatrix} 0 & 1 & 0 \\ 0 & -t & 1 \\ 1 & 0 & -t \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -t & 1 \end{vmatrix} = 1.$$

It follows from (*) that

$$p(t) = (-t)(-t)^3 - 1 = t^4 - 1.$$

The eigenvalues of $A$ are the roots of the characteristic polynomial $p(t)$.

Solving $t^4 - 1 = 0$, we obtain the eigenvalues

$$\pm 1, \ \pm i,$$

where $i = \sqrt{-1}$.

Note that $t^4 - 1 = (t-1)(t+1)(t-i)(t+i)$.

16 Let $A$ be an $n \times n$ matrix. We say that $A$ is **idempotent** if $A^2 = A$.
**(a)** Find a nonzero, nonidentity idempotent matrix.
**(b)** Show that eigenvalues of an idempotent matrix $A$ is either $0$ or $1$.
(*The Ohio State University*)

## Proof.

### (a) Nonzero, nonidentity idempotent matrix

Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. Then $A$ is a nonzero, nonidentity matrix and $A$ is idempotent since we have

$$A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = A.$$

### (b) Eigenvalues of an idempotent matrix $A$ is either 0 or 1

Let $\lambda$ be an eigenvalue of the idempotent matrix $A$ and let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$.

Namely we have

$$A\mathbf{x} = \lambda\mathbf{x}, \mathbf{x} \neq \mathbf{0}.$$

Then we compute $A^2\mathbf{x}$ in two ways.

First, since $A$ is idempotent we have $A^2 = A$ and thus we have

$$A^2\mathbf{x} = A\mathbf{x} \overset{(*)}{=} \lambda\mathbf{x}.$$

---

Next, we compute as follows.

$$A^2\mathbf{x} = A(A\mathbf{x}) \overset{(*)}{=} A(\lambda\mathbf{x}) = \lambda(A\mathbf{x}) \overset{(*)}{=} \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}.$$

Comparing these two computations, we obtain

$$\lambda\mathbf{x} = \lambda^2\mathbf{x}.$$

Since $\mathbf{x}$ is a nonzero vector (because $\mathbf{x}$ is an eigenvector), we must have

$$\lambda = \lambda^2.$$

Hence solving $\lambda(\lambda - 1) = 0$, the possible values for $\lambda$ is either 0 or 1.

Thus, the idempotent matrix $A$ only have eigenvalues 0 or 1.

---

17 A square matrix $A$ is called **idempotent** if $A^2 = A$.
**(a)** Let $\mathbf{u}$ be a vector in $\mathbb{R}^n$ with length $1$.
Define the matrix $P$ to be $P = \mathbf{u}\mathbf{u}^{\mathrm{T}}$. Prove that $P$ is an idempotent matrix.
**(b)** Suppose that $\mathbf{u}$ and $\mathbf{v}$ be unit vectors in $\mathbb{R}^n$ such that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.
Let $Q = \mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}$. Prove that $Q$ is an idempotent matrix.
**(c)** Prove that each nonzero vector of the form $a\mathbf{u} + b\mathbf{v}$ for some $a, b \in \mathbb{R}$ is an eigenvector corresponding to the eigenvalue $1$ for the matrix $Q$ in part (b).

## Proof.

### (a) Prove that $P = \mathbf{u}\mathbf{u}^{\mathrm{T}}$ is an idempotent matrix.

The length of the vector $\mathbf{u}$ is given by

$$\|\mathbf{u}\| = \sqrt{\mathbf{u}^{\mathrm{T}}\mathbf{u}}.$$

Since $\|\mathbf{u}\| = 1$ by assumption, it yields that

$$\mathbf{u}^{\mathrm{T}}\mathbf{u} = 1.$$

Let us compute $P^2$ using the associative properties of matrix multiplication.

We have

$$\begin{aligned} P^2 &= (\mathbf{u}\mathbf{u}^{\mathrm{T}})(\mathbf{u}\mathbf{u}^{\mathrm{T}}) \\ &= \mathbf{u}(\mathbf{u}^{\mathrm{T}}\mathbf{u})\mathbf{u}^{\mathrm{T}} \\ &= \mathbf{u}(1)\mathbf{u}^{\mathrm{T}} = \mathbf{u}\mathbf{u}^{\mathrm{T}} = P. \end{aligned}$$

Thus, we have obtained $P^2 = P$, and hence $P$ is an idempotent matrix.

### (b) Prove that $Q = \mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}$ is an idempotent matrix

Since **u** and **v** are unit vectors, we have as in part (a)

$$\mathbf{u}^{\mathrm{T}}\mathbf{u} = 1 \text{ and } \mathbf{v}^{\mathrm{T}}\mathbf{v} = 1.$$

Since **u** and **v** are orthogonal, their dot (inner) product is 0.

Thus we have

$$\mathbf{u}^{\mathrm{T}}\mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{u} = 0.$$

Using these identities, we compute $Q^2$ as follows.

We have

$$\begin{aligned} Q^2 &= (\mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}})(\mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}) \\ &= \mathbf{u}\mathbf{u}^{\mathrm{T}}(\mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}) + \mathbf{v}\mathbf{v}^{\mathrm{T}}(\mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}}) \\ &= \mathbf{u}\underbrace{\mathbf{u}^{\mathrm{T}}\mathbf{u}}_{1}\mathbf{u}^{\mathrm{T}} + \mathbf{u}\underbrace{\mathbf{u}^{\mathrm{T}}\mathbf{v}}_{0}\mathbf{v}^{\mathrm{T}} + \mathbf{v}\underbrace{\mathbf{v}^{\mathrm{T}}\mathbf{u}}_{0}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\underbrace{\mathbf{v}^{\mathrm{T}}\mathbf{v}}_{1}\mathbf{v}^{\mathrm{T}} \\ &= \mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}} = Q. \end{aligned}$$

It follows that $Q$ is an idempotent matrix.

### (c) Prove that $a\mathbf{u} + b\mathbf{v}$ is an eigenvector

Let us first compute $Q\mathbf{u}$. We have

$$\begin{aligned} Q\mathbf{u} &= (\mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{v}\mathbf{v}^{\mathrm{T}})\mathbf{u} \\ &= \mathbf{u}\underbrace{\mathbf{u}^{\mathrm{T}}\mathbf{u}}_{1} + \mathbf{v}\underbrace{\mathbf{v}^{\mathrm{T}}\mathbf{u}}_{0} = \mathbf{u}. \end{aligned}$$

Note that **u** is a nonzero vector because it is a unit vector.

Thus, the equality $Q\mathbf{u} = \mathbf{u}$ implies that 1 is an eigenvalue of $Q$ and **v** is a corresponding eigenvector.

Similarly, we can check that **v** is an eigenvector corresponding to the eigenvalue 1.

Now let $a\mathbf{u} + b\mathbf{v}$ be a nonzero vector.

Then we have

$$Q(a\mathbf{u} + b\mathbf{v}) = aQ\mathbf{u} + bQ\mathbf{v} = a\mathbf{u} + b\mathbf{v}.$$

It follows that $a\mathbf{u} + b\mathbf{v}$ is an eigenvector corresponding to the eigenvalue 1.

Another way to prove (c)

Another way to see this is as follows.

As we saw above, the vectors $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors corresponding to the eigenvalue $1$.

Hence $\mathbf{u}, \mathbf{v} \in E_1$, where $E_1$ is an eigenspace of the eigenvalue 1

Note that $E_1$ is a vector space, hence $a\mathbf{u} + b\mathbf{v}$ is a nonzero vector in $E_1$.

Thus, $a\mathbf{u} + b\mathbf{v}$ is an eigenvector corresponding to the eigenvalue $1$ as well.

## Related Question.

**Problem.**

**(a)** Find a nonzero, nonidentity idempotent matrix.

**(b)** Show that eigenvalues of an idempotent matrix $A$ is either 0 or 1.

---

18 Let $A$ be an $n \times n$ matrix. Suppose that the matrix $A^2$ has a real eigenvalue $\lambda > 0$. Then show that either $\sqrt{\lambda}$ or $-\sqrt{\lambda}$ is an eigenvalue of the matrix $A$.

## Hint.

Use the following fact: a scalar $\lambda$ is an eigenvalue of a matrix $A$ if and only if

$$\det(A - \lambda I) = 0.$$

## Proof.

Since $\lambda$ is an eigenvalue of $A^2$, the determinant of the matrix $A^2 - \lambda I$ is zero, where $I$ is the $n \times n$ identity matrix:

$$\det(A^2 - \lambda I) = 0.$$

Now we have the following factorization.

$$A^2 - \lambda I = (A - \sqrt{\lambda} I)(A + \sqrt{\lambda} I).$$

Taking the determinant of both sides, we obtain

$$\begin{aligned} 0 &= \det(A^2 - \lambda I) = \det\big((A - \sqrt{\lambda} I)(A + \sqrt{\lambda} I)\big) \\ &= \det(A - \sqrt{\lambda} I)\det(A + \sqrt{\lambda} I) \end{aligned}$$

by the multiplicative property of the determinant.

Therefore we have either

$$\det(A - \sqrt{\lambda} I) = 0 \text{ or } \det(A + \sqrt{\lambda} I) = 0$$

Thus we conclude that either $\sqrt{\lambda}$ or $-\sqrt{\lambda}$ is an eigenvalue of the matrix $A$.

**19** Show that if $A$ and $B$ are similar matrices, then they have the same eigenvalues and their algebraic multiplicities are the same.

## Proof.

We prove that $A$ and $B$ have the same characteristic polynomial. Then the result follows immediately since eigenvalues and algebraic multiplicities of a matrix are determined by its characteristic polynomial.

Since $A$ and $B$ are similar, there exists an invertible matrix $S$ such that $S^{-1}AS = B$.
Let $p_A(t)$ and $p_B(t)$ denote the characteristic polynomials of $A$ and $B$, respectively.

We have

$$\begin{aligned} p_B(t) &= \det(B - tI) = \det(S^{-1}AS - tI) \\ &= \det(S^{-1}(A - tI)S) = \det(S^{-1})\det(A - tI)\det(S) \\ &\overset{(*)}{=} \det(A - tI) = p_A(t). \end{aligned}$$

Here the fifth equality (*) follows from the fact that $\det(S^{-1}) = \det(S)^{-1}$.
(Also note that even though matrix multiplication is not commutative in general but determinants are just numbers, thus we can change the order of the product of determinants.)

Thus we showed that $p_A(t) = p_B(t)$ and this completes the proof.

---

**20** For a real number $0 \leq \theta \leq \pi$, we define the real $3 \times 3$ matrix $A$ by

$$A = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**(a)** Find the determinant of the matrix $A$.
**(b)** Show that $A$ is an orthogonal matrix.
**(c)** Find the eigenvalues of $A$.

## Solution.

### (a) The determinant of the matrix $A$

By the cofactor expansion corresponding to the third row, we compute

$$\begin{aligned} \det(A) &= \begin{vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} \\ &= 0 \cdot \begin{vmatrix} -\sin\theta & 0 \\ \cos\theta & 0 \end{vmatrix} - 0 \cdot \begin{vmatrix} \cos\theta & 0 \\ \sin\theta & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} \\ &= \cos^2\theta + \sin^2\theta \\ &= 1. \end{aligned}$$

The last step follows from the famous trigonometry identity

$$\cos^2\theta + \sin^2\theta = 1.$$

Thus we have

$$\det(A) = 1.$$

### (b) The matrix $A$ is an orthogonal matrix

We give two solutions for part (b).

#### The first solution of (b)

The first solution computes $A^{\mathrm{T}}A$ and show that it is the identity matrix $I$.

We have

$$\begin{aligned} A^{\mathrm{T}}A &= \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 & 0 \\ 0 & \cos^2\theta + \sin^2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I. \end{aligned}$$

Similarly, you can check that $AA^{\mathrm{T}} = I$. Thus $A$ is an orthogonal matrix.

#### The second solution of (b)

The second proof uses the following fact: a matrix is orthogonal if and only its column vectors form an orthonormal set.

Let

$$A_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, A_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

be the column vectors of the matrix $A$. The length of these vectors are all $1$. For example, we have

$$||A_1|| = \sqrt{(\cos\theta)^2 + (\sin\theta)^2 + 0^2} = \sqrt{1} = 1.$$

Similarly, we have $||A_2|| = ||A_3|| = 1$.

The dot (inner) product of $A_1$ and $A_2$ is

$$A_1 \cdot A_2 = \cos\theta \cdot (-\sin\theta) + \sin\theta \cdot \cos\theta + 0 \cdot 0 = 0.$$

Similarly, we have $A_1 \cdot A_3 = A_2 \cdot A_3 = 0$.

Therefore, the column vectors $A_1, A_2, A_3$ are orthonormal vectors. Hence by the above fact, the matrix $A$ is orthogonal.

## (c) The eigenvalues of $A$

We compute the characteristic polynomial $p(t)=\det(A-tI)$ as follows.

$$\begin{aligned}
p(t)=\det(A-tI)&=\begin{vmatrix} \cos\theta-t & -\sin\theta & 0 \\ \sin\theta & \cos\theta-t & 0 \\ 0 & 0 & 1-t \end{vmatrix}\\
&=(1-t)\begin{vmatrix} \cos\theta-t & -\sin\theta \\ \sin\theta & \cos\theta-t \end{vmatrix} \text{ by the third row cofactor expansion}\\
&=(1-t)(\cos^2\theta-2t\cos\theta+t^2+\sin^2\theta)\\
&=(1-t)(t^2-(2\cos\theta)t+1).
\end{aligned}$$

The eigenvalues are roots of the characteristic polynomial $p(t)$, hence we solve

$$p(t)=(1-t)(t^2-(2\cos\theta)t+1)=0.$$

One solution is $t=1$. The other solutions come from the quadratic polynomial in $p(t)$.
By the quadratic formula, those solutions are

$$\begin{aligned}
t&=\cos\theta\pm\sqrt{\cos^2\theta-1}\\
&=\cos\theta\pm\sqrt{-\sin^2\theta}\\
&=\cos\theta\pm i\sin\theta
\end{aligned}$$

since $\sin\theta\geq 0$ since $0\leq\theta\leq\pi$.
Therefore the eigenvalues of the matrix $A$ are

$$1, \cos\theta\pm i\sin\theta.$$

## Related Question.

The following problem treats the rotation matrix in the plane.

**Problem**.
Consider the $2\times 2$ matrix

$$A=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

where $\theta$ is a real number $0\leq\theta<2\pi$.

**(a)** Find the characteristic polynomial of the matrix $A$.

**(b)** Find the eigenvalues of the matrix $A$.

**(c)** Determine the eigenvectors corresponding to each of the eigenvalues of $A$.

㉑ Let $A$ be an $n \times n$ matrix. Assume that every vector $\mathbf{x}$ in $\mathbb{R}^n$ is an eigenvector for some eigenvalue of $A$. Prove that there exists $\lambda \in \mathbb{R}$ such that $A = \lambda I$, where $I$ is the $n \times n$ identity matrix.

---

## Proof.

Let us write $A = (a_{ij})$.

Let $\mathrm{e}_i$ be the unit vector in $\mathbb{R}^n$ whose $i$-th entry is $1$, and $0$ elsewhere.

Then the vector $\mathrm{e}_i$ is an eigenvector corresponding to some eigenvalue $\lambda_i$. That is, we have

$$A\mathrm{e}_i = \lambda_i \mathrm{e}_i.$$

More explicitly, this shows that

$$\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{n-1i} \\ a_{ni} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{bmatrix}.$$

Therefore, it follows that $a_{ki} = 0$ if $k \neq i$ and $a_{ii} = \lambda_i$.

So we have

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Now we consider the vector $\mathrm{v} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$ with all entries $1$. Then $\mathrm{v}$ is an eigenvector of an eigenvalue $\lambda$, that is, $A\mathrm{v} = \lambda \mathrm{v}$.

It follows that we have

$$\begin{bmatrix} \lambda \\ \lambda \\ \vdots \\ \lambda \end{bmatrix} = \lambda \mathrm{v} = A\mathrm{v} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

As a result, we must have $\lambda = \lambda_1 = \cdots = \lambda_n$.

Thus, we obtain

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{bmatrix} = \lambda I$$

as required.

22 Let $A$ be a square matrix. Prove that the eigenvalues of the transpose $A^{\mathrm{T}}$ are the same as the eigenvalues of $A$.

---

### Proof.

Recall that the eigenvalues of a matrix are roots of its characteristic polynomial.
Hence if the matrices $A$ and $A^{\mathrm{T}}$ have the same characteristic polynomial, then they have the same eigenvalues.

So we show that the characteristic polynomial $p_A(t) = \det(A - tI)$ of $A$ is the same as the characteristic polynomial $p_{A^{\mathrm{T}}}(t) = \det(A^{\mathrm{T}} - tI)$ of the transpose $A^{\mathrm{T}}$.

We have

$$\begin{aligned}
p_{A^{\mathrm{T}}}(t) &= \det(A^{\mathrm{T}} - tI) \\
&= \det(A^{\mathrm{T}} - tI^{\mathrm{T}}) && \text{since } I^{\mathrm{T}} = I \\
&= \det\left((A - tI)^{\mathrm{T}}\right) \\
&= \det(A - tI) && \text{since } \det(B^{\mathrm{T}}) = \det(B) \text{ for any square matrix } B \\
&= p_A(t).
\end{aligned}$$

Therefore we obtain $p_{A^{\mathrm{T}}}(t) = p_A(t)$, and we conclude that the eigenvalues of $A$ and $A^{\mathrm{T}}$ are the same.

### Remark: Algebraic Multiplicities of Eigenvalues

Remark that since the characteristic polynomials of $A$ and the transpose $A^{\mathrm{T}}$ are the same, it furthermore yields that the algebraic multiplicities of eigenvalues of $A$ and $A^{\mathrm{T}}$ are the same.

23 Let $A$ be an $n$ by $n$ matrix with entries in complex numbers $\mathbb{C}$. Its only eigenvalues are $1, 2, 3, 4, 5$, possibly with multiplicities. What is the rank of the matrix $A + I_n$, where $I_n$ is the identity $n$ by $n$ matrix.
(*University of California, Berkeley*)

---

### Hint.

The problem is asking if you understand the definition of eigenvalues and basic relation between the rank and non-singularlity.

### Solution.

The matrix $A + I_n$ is nonsingular, otherwise $-1$ is an eigenvalue of $A$ but by assumption it is impossible.
Since the matrix $A + I_n$ is nonsingular, it has full rank. Since $A + I_n$ is $n$ by $n$ matrix, its rank must be $n$.

### Comment.

The solution is very short and simple. The point is to notice that $A + I_n$ is of the familiar form $A - \lambda I_n$.

If the problem asked to find the rank of $A - 6I_n$, then it would have been easier to notice this.

### Related Question.

A related question is:

What is the nullity of $A$ ?

See the post "Find the nullity of the matrix $A+I$ if eigenvalues are $1, 2, 3, 4, 5$" for a solution.

24 Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an $2\times 2$ matrix. Express the eigenvalues of $A$ in terms of the trace and the determinant of $A$.

## Solution.

Recall the definitions of the trace and determinant of $A$ :

$$\operatorname{tr}(A)=a+d \text{ and } \det(A)=ad-bc.$$

---

The eigenvalues of $A$ are roots of the characteristic polynomial $p(t)$ of $A$ . So let us first find $p(t)$ .

We have

$$\begin{aligned} p(t)=\det(A-tI)&=\begin{vmatrix} a-t & b \\ c & d-t \end{vmatrix} \\ &=(a-t)(d-t)-bc \\ &=t^2-(a+d)t+ad-bc \\ &=t^2-\operatorname{tr}(A)t+\det(A). \end{aligned}$$

Using the quadratic formula, the eigenvalues of A (roots of $p(t)$ ) are

$$\frac{\operatorname{tr}(A)\pm\sqrt{\operatorname{tr}(A)^2-4\det(A)}}{2}.$$

25 Let $A$ be an $n\times n$ nilpotent matrix, that is, $A^m=O$ for some positive integer $m$, where $O$ is the $n\times n$ zero matrix.

Prove that $A$ is a singular matrix and also prove that $I-A, I+A$ are both nonsingular matrices, where $I$ is the $n\times n$ identity matrix.

## Hint.

We give two proofs. The first one uses only some properties of the determinant of a matrix.

The second one uses the theory of eigenvalues.

We summarize several facts we will use below.

---

In proof 1.

26 A matrix is nonsingular if and only if it is invertible.

㉗ A matrix is nonsingular if and only if its determinant is nonzero.

㉘ For two square matrices $A, B$, we have

$$\det(AB) = \det(A)\det(B).$$

In proof 2.

㉙ Eigenvalues of $A$ are all scalars $\lambda$ such that the matrix $A - \lambda I$ is nonsingular.

## Proof 1.

We use the fact that a matrix is nonsingular if and only if its determinant is nonzero.

Since the determinant is multiplicative, we have

$$0 = \det(O) = \det(A^m) = \det(A)^m.$$

This implies that $\det(A) = 0$, and hence the matrix $A$ is singular.

Next, we show that the matrix $I - A$ is nonsingular. Note that we have

$$I = I - A^m = (I - A)(I + A + A^2 + \cdots + A^{m-1}).$$

This implies that the matrix $I + A + A^2 + \cdots + A^{m-1}$ is the inverse matrix of $I - A$.

Thus the matrix $I - A$ is invertible, and hence it is nonsingular.

Next, we show that the matrix $I + A$ is nonsingular.

Note that we have

$$I - (-A)^m = (I + A)(I - A + A^2 - \cdots + (-1)^{m-1}A^{m-1}).$$

Since $(-A)^m = (-1)^m A^m = (-1)^m O = O$, we obtain

$$I = (I + A)(I - A + A^2 - \cdots + (-1)^{m-1}A^{m-1})$$

and this implies that $I + A$ is invertible matrix with the inverse matrix

$$(I + A)^{-1} = I - A + A^2 - \cdots + (-1)^{m-1}A^{m-1}.$$

Therefore, the matrix $I + A$ is nonsingular.

## Proof 2.

In the second proof, we will use an eigenvalue technique.

Let $\lambda$ be an eigenvalue of the matrix $A$ and let $\mathbf{x}$ be the corresponding eigenvector. Namely, we have

$$A\mathrm{x} = \lambda\mathrm{x}.$$

Then we have

$$\begin{aligned} 0 &= O\mathrm{x} = A^m\mathrm{x} \\ &= A^{m-1}A\mathrm{x} = A^{m-1}(\lambda\mathrm{x}) = \cdots = \lambda^m\mathrm{x}. \end{aligned}$$

It follows that $\lambda = 0$. Thus the only eigenvalue of $A$ is zero.

Recall that eigenvalues $\lambda$ are all scalars such that $A - \lambda I$ is a singular matrix.
Since $\lambda = 0$ is an eigenvalue of $A$, the matrix $A = A - 0I$ is singular.

Since $1$ is not an eigenvalue, we see that $A - I$ is nonsingular, and hence $I - A = -(A - I)$ is also nonsingular.

Similarly, since $-1$ is not an eigenvalue, we observe that $I + A = A + I$ is nonsingular.

## Related Question.

A natural source of nilpotent matrices is the differentiation linear transformation of vector space of polynomials.

Check out the post "Differentiating Linear Transformation is Nilpotent" for problems and solutions about the nilpotency of the differentiation linear transformation.

## Related Question.

Here is another problem about a nilpotent matrix.

**Problem**.
Suppose that $A$ is an $n \times n$ nilpotent matrix and $B$ is an $n \times n$ invertible matrix.
Is the matrix $B - A$ invertible? If so, give a proof. Otherwise, give a counterexample.

㉚ Determine whether each of the following statements is True or False.
**(a)** If $A$ and $B$ are $n \times n$ matrices, and $P$ is an invertible $n \times n$ matrix such that $A = PBP^{-1}$, then $\det(A) = \det(B)$.
**(b)** If the characteristic polynomial of an $n \times n$ matrix $A$ is $p(\lambda) = (\lambda - 1)^n + 2$, then $A$ is invertible.
**(c)** If $A^2$ is an invertible $n \times n$ matrix, then $A^3$ is also invertible.
**(d)** If $A$ is a $3 \times 3$ matrix such that $\det(A) = 7$, then $\det(2A^{\mathrm{T}}A^{-1}) = 2$.

**(e)** If $\mathbf{v}$ is an eigenvector of an $n \times n$ matrix $A$ with corresponding eigenvalue $\lambda_1$, and if $\mathbf{w}$ is an eigenvector of $A$ with corresponding eigenvalue $\lambda_2$, then $\mathbf{v} + \mathbf{w}$ is an eigenvector of $A$ with corresponding eigenvalue $\lambda_1 + \lambda_2$.
(*Stanford University, Linear Algebra Exam Problem*)

## Solution.

We use the following properties of the determinant.

㉛ For $n \times n$ matrices $A, B$, we have

$$\det(AB) = \det(A)\det(B).$$

---

㉜ If $A$ is invertible, we have

$$\det(A^{-1}) = \det(A)^{-1}.$$

---

Also recall that a matrix $A$ is invertible if and only if $\det(A) \neq 0$.

(a) True or False: If $A = PBP^{-1}$, then $\det(A) = \det(B)$.

True. We have

$$\begin{aligned} \det(A) &= \det(PBP^{-1}) \\ &= \det(P)\det(B)\det(P)^{-1} && \text{by properties 1, 2} \\ &= \det(P)\det(P)^{-1}\det(B) && \text{since determinants are just numbers} \\ &= \det(B). \end{aligned}$$

(b) True or False: If $p(\lambda) = (\lambda - 1)^n + 2$, then $A$ is invertible.

True. Note that we have $p(0) = (-1)^n + 2 \neq 0$.

Thus $0$ is not an eigenvalue of $A$.

Recall that a matrix is invertible if and only if it does not have $0$ as an eigenvalue.

Thus, $A$ is invertible.

(c) True or False: If $A^2$ is an invertible $n \times n$ matrix, then $A^3$ is also invertible.

True. If $A^2$ is invertible, then we have $\det(A^2) \neq 0$.

Then by property 1, we have

$$\det(A)^2 = \det(A^2) \neq 0,$$

and hence $\det(A) \neq 0$.

It follows from property 1 that we have

$$\det(A^3) = \det(A)^3 \neq 0.$$

Thus the matrix $A^3$ is invertible.

(d) True or False: If $A$ is a $3 \times 3$ matrix such that $\det(A) = 7$, then $\det(2A^{\mathrm{T}}A^{-1}) = 2$.

False. Recall that we have $\det(A^{\mathrm{T}}) = \det(A)$.

Together with properties 1, 2, we have

$$\begin{aligned}\det(2A^{\mathrm{T}}A^{-1}) &= \det(2I)\det(A^{\mathrm{T}})\det(A^{-1})\\ &= \det(2I)\det(A)\det(A)^{-1}\\ &= \det(2I) = 8.\end{aligned}$$

Here $I$ is the $3 \times 3$ identity matrix.

Remark that in general for an $n \times n$ matrix and a scalar $r$, we have

$$\det(rA) = r^n \det(A).$$

> (e) True or False: $\mathbf{v} + \mathbf{w}$ is an eigenvector of $A$ with corresponding eigenvalue $\lambda_1 + \lambda_2$.

False. For example, let $A$ be a $2 \times 2$ identity matrix.
The only eigenvalue of $A$ is $1$.
Note that $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are both eigenvector corresponding to the eigenvalue $\lambda_1 = \lambda_2 = 1$.
However, $\lambda_1 + \lambda_2 = 2$ is not an eigenvalue of $A$, and thus $\mathbf{v} + \mathbf{w}$ cannot be an eigenvector corresponding to $\lambda_1 + \lambda_2 = 2$.

Note that the set of eigenvectors corresponding to an eigenvalue $\lambda$ together with the zero vector form a vector space, called the eigenspace of $\lambda$. Thus, if $\mathbf{v}$ and $\mathbf{w}$ are both eigenvectors of $\lambda$, then the sum $\mathbf{v} + \mathbf{w}$ is also an eigenvector of $\lambda$ or the zero vector.

---

33 Let $A = (a_{ij})$ be an $n \times n$ matrix. We say that $A = (a_{ij})$ is a **right stochastic matrix** if each entry $a_{ij}$ is nonnegative and the sum of the entries of each row is $1$. That is, we have

$$a_{ij} \geq 0 \quad \text{and} \quad a_{i1} + a_{i2} + \cdots + a_{in} = 1$$

for $1 \leq i, j \leq n$. Let $A = (a_{ij})$ be an $n \times n$ right stochastic matrix. Then show the following statements.
**(a)**The stochastic matrix $A$ has an eigenvalue $1$.
**(b)** The absolute value of any eigenvalue of the stochastic matrix $A$ is less than or equal to $1$.

---

## Proof.

> (a) The stochastic matrix $A$ has an eigenvalue $1$.

We compute that

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} + a_{12} + \cdots + a_{1n} \\ a_{21} + a_{22} + \cdots + a_{2n} \\ \vdots \\ a_{n1} + a_{n2} \cdots + a_{nn} \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Here the second equality follows from the definition of a right stochastic matrix.
(Each row sums up to $1$.)

This computation shows that $1$ is an eigenvector of $A$ and $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $1$.

> (b) The absolute value of any eigenvalue of the stochastic matrix $A$ is less than or equal to $1$.

Let $\lambda$ be an eigenvalue of the stochastic matrix $A$ and let $\mathrm{v}$ be a corresponding eigenvector.
That is, we have

$$A\mathrm{v} = \lambda \mathrm{v}.$$

Comparing the $i$-th row of the both sides, we obtain

$$a_{i1} v_1 + a_{i2} v_2 + \cdots + a_{in} v_n = \lambda v_i \qquad (*)$$

for $i = 1, \dots, n$.
Let

$$|v_k| = \max\{|v_1|, |v_2|, \dots, |v_n|\},$$

namely $v_k$ is the entry of $\mathrm{v}$ that has the maximal absolute value.

Note that $|v_k| > 0$ since otherwise we have $\mathrm{v} = 0$ and this contradicts that an eigenvector is a nonzero vector.
Then from (*) with $i = k$, we have

$$\begin{aligned} |\lambda| \cdot |v_k| &= |a_{k1} v_1 + a_{k2} v_2 + \cdots + a_{kn} v_n| \\ &\leq a_{k1} |v_1| + a_2 |v_2| + \cdots + a_{kn} |v_n| && \text{(by the triangle inequality and } a_{ij} \geq 0 \\ &\leq a_{k1} |v_k| + a_2 |v_k| + \cdots + a_{kn} |v_k| && \text{(since } |v_k| \text{ is maximal)} \\ &= (a_{k1} + a_{k2} + \cdots + a_{kn})|v_k| = |v_k|. \end{aligned}$$

Since $|v_k| > 0$, it follows that

$$\lambda \leq 1$$

as required.

## Remark.

A stochastic matrix is also called probability matrix, transition matrix, substitution matrix, or Markov matrix.

㉞ Suppose that $\lambda$ and $\mu$ are two distinct eigenvalues of a square matrix $A$ and let $\mathbf{x}$ and $\mathbf{y}$ be eigenvectors corresponding to $\lambda$ and $\mu$, respectively. If $a$ and $b$ are nonzero numbers, then prove that $a\mathbf{x} + b\mathbf{y}$ is not an eigenvector of $A$ (corresponding to any eigenvalue of $A$).

## Hint.

We use the following fact in the proof.

**Fact**: Two eigenvectors corresponding to distinct eigenvalues are linearly independent.

### Proof.

Seeking a contradiction, we assume that $a\mathbf{x} + b\mathbf{y}$ is an eigenvector corresponding to an eigenvalue $\zeta$. Thus we have

$$A(a\mathbf{x} + b\mathbf{y}) = \zeta(a\mathbf{x} + b\mathbf{y}). \tag{*}$$

---

We calculate the left hand side of this equality as follows.

We have

$$\begin{aligned} A(a\mathbf{x} + b\mathbf{y}) &= aA\mathbf{x} + bA\mathbf{y} \\ &= a\lambda\mathbf{x} + b\mu\mathbf{y} \end{aligned}$$

since $A\mathbf{x} = \lambda\mathbf{x}, A\mathbf{y} = \mu\mathbf{y}$ by defintion.

Therefore, from (*) we obtain

$$a(\lambda - \zeta)\mathbf{x} + b(\mu - \zeta)\mathbf{y} = 0.$$

---

Recall that eigenvectors corresponding to distinct eigenvalues are linearly independent. Thus $\mathbf{x}$ and $\mathbf{y}$ are linearly independent.

Thus, the coefficients of the above linear combinations must be zero:

$$a(\lambda - \zeta) = 0 \text{ and } b(\mu - \zeta) = 0.$$

Since $a \neq 0, b \neq 0$, this implies that we have

$$\lambda = \zeta = \mu,$$

and this is a contradiction because $\lambda$ and $\mu$ are supposed to be distinct.

Hence, $a\mathbf{x} + b\mathbf{y}$ cannot be an eigenvector of any eigenvalue of $A$.

(35) Suppose that $A$ is an $n \times n$ singular matrix. Prove that for sufficiently small $\epsilon > 0$, the matrix $A - \epsilon I$ is nonsingular, where $I$ is the $n \times n$ identity matrix.

### Hint.

Consider the characteristic polynomial $p(t)$ of the matrix $A$.

Note that the eigenvalues of $A$ are the roots of $p(t)$.

Thus if $\epsilon$ is not an eigenvalue, then $p(\epsilon) \neq 0$.

### Proof.

Let $p(t) = \prod_{i=}^{n} (\lambda_i - t)$ be the characteristic polynomial for $A$, where $\lambda_i$ are eigenvalues of $A$.

Let $\lambda_{i_0}$ be the nonzero eigenvalue of $A$ of the smallest absolute value.
That is $|\lambda_{i_0}| \leq |\lambda_i|$ for any nonzero eigenvalue $\lambda_i$.
Then for any $0 < \epsilon < |\lambda_{i_0}|$, we have $\det(A - \epsilon I) = p(\epsilon) \neq 0$, otherwise $\epsilon$ would be an eigenvalue of $A$ but it is impossible because of the minimality of $\lambda_{i_0}$.

Therefore the matrix $A - \epsilon I$ is nonsingular for all $0 < \epsilon < |\lambda_{i_0}|$.

(36) Let $A$ and $B$ be $n \times n$ matrices. Prove that the characteristic polynomials for the matrices $AB$ and $BA$ are the same.

## Hint.

(37) Consider the case when the matrix $A$ is invertible.

(38) Even if $A$ is not invertible, note that $A - \epsilon I$ is invertible matrix for sufficiently small $\epsilon$. (See Problem Perturbation of a singular matrix is nonsingular for a proof of this fact.)

(39) Take the limit $\epsilon \to 0$.

## Proof.

We want to show that $|AB - \lambda I| = |BA - \lambda I|$, where $\lambda$ is an unknown and $I$ is the $n \times n$ identity matrix.

First let us consider the case when the matrix $A$ is invertible.
So suppose that $A$ is invertible.
Then we have

$$|AB - \lambda I| = |A^{-1}(AB - \lambda I)A| = |BA - \lambda I|.$$

Here we use the fact that $|A||A^{-1}| = |AA^{-1}| = |I| = 1$.
Thus when $A$ is invertible, the claim is proved.

Now we consider the general case.
Whether or not $A$ is invertible, the matrix $A - \epsilon I$ is invertible for sufficiently small $\epsilon$.

(More precisely, if $\epsilon$ is smaller than absolute values of all nonzero eigenvalues of $A$. then $A - \epsilon I$ is invertible. See Problem Perturbation of a singular matrix is nonsingular for a proof.)

Then by the previous case, we have

$$|(A-\epsilon I)B-\lambda I| = |B(A-\epsilon I)-\lambda I|.$$

Taking the limit $\epsilon \to 0$, we obtain

$$|AB-\lambda I| = |BA-\lambda I|.$$

### Comment.

When you study hard linear algebra, you might have forgotten about calculus.
This problem suggests that for some problems in linear algebra, techniques from calculus (like limits) might help to solve them.

(40) Let $A$ be an $n \times n$ real matrix. Prove that if $\lambda$ is an eigenvalue of $A$, then its complex conjugate $\bar{\lambda}$ is also an eigenvalue of $A$.

### Proof 1.

Let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$. Then we have

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Taking the conjugate of both sides, we have

$$\overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}}.$$

Since $A$ is a real matrix, it yields that

$$A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}. \qquad (*)$$

Note that $\mathbf{x}$ is a nonzero vector as it is an eigenvector. Then the complex conjugate $\bar{\mathbf{x}}$ is a nonzero vector as well.
Thus the equality (*) implies that the complex conjugate $\bar{\lambda}$ is an eigenvalue of $A$ with corresponding eigenvector $\bar{\mathbf{x}}$.

### Proof 2.

Let $p(t)$ be the characteristic polynomial of $A$.
Recall that the roots of the characteristic polynomial $p(t)$ are the eigenvalues of $A$.
Thus, we have

$$p(\lambda) = 0.$$

As $A$ is a real matrix, the characteristic polynomial $p(t)$ has real coefficients.
It follows that

$$\overline{p(t)} = p(\,\bar{t}\,).$$

The previous two identities yield that

$$p(\bar{\lambda}) = \overline{p(\lambda)} = \bar{0} = 0,$$

and the complex conjugate $\bar{\lambda}$ is a root of $p(t)$, and hence $\bar{\lambda}$ is an eigenvalue of $A$.

㊶ Let $A$ be a square matrix. Prove that the eigenvalues of the transpose $A^{\mathrm{T}}$ are the same as the eigenvalues of $A$.

## Proof.

Recall that the eigenvalues of a matrix are roots of its characteristic polynomial.
Hence if the matrices $A$ and $A^{\mathrm{T}}$ have the same characteristic polynomial, then they have the same eigenvalues.

So we show that the characteristic polynomial $p_A(t) = \det(A - tI)$ of $A$ is the same as the characteristic polynomial $p_{A^{\mathrm{T}}}(t) = \det(A^{\mathrm{T}} - tI)$ of the transpose $A^{\mathrm{T}}$.

We have

$$\begin{aligned}
p_{A^{\mathrm{T}}}(t) &= \det(A^{\mathrm{T}} - tI) \\
&= \det(A^{\mathrm{T}} - tI^{\mathrm{T}}) && \text{since } I^{\mathrm{T}} = I \\
&= \det\left((A - tI)^{\mathrm{T}}\right) \\
&= \det(A - tI) && \text{since } \det(B^{\mathrm{T}}) = \det(B) \text{ for any square matrix } B \\
&= p_A(t).
\end{aligned}$$

Therefore we obtain $p_{A^{\mathrm{T}}}(t) = p_A(t)$, and we conclude that the eigenvalues of $A$ and $A^{\mathrm{T}}$ are the same.

### Remark: Algebraic Multiplicities of Eigenvalues

Remark that since the characteristic polynomials of $A$ and the transpose $A^{\mathrm{T}}$ are the same, it furthermore yields that the algebraic multiplicities of eigenvalues of $A$ and $A^{\mathrm{T}}$ are the same.

㊷ Let $A$ be an $n \times n$ matrix. Its only eigenvalues are $1, 2, 3, 4, 5$, possibly with multiplicities. What is the nullity of the matrix $A + I_n$, where $I_n$ is the $n \times n$ identity matrix?
(*The Ohio State University*)

## Solution.

Recall that $\lambda$ is an eigenvalue of the matrix $A$ if and only if the matrix $A - \lambda I$ is singular.
Thus, $A + I$ is nonsingular, otherwise $-1$ is an eigenvalue of $A$.

Since $A + I$ is nonsingular, the we have the null space $\mathcal{N}(A + I) = \{0\}$, and hence the nullity of $A + I$ is zero.

㊸ Suppose that a real symmetric matrix $A$ has two distinct eigenvalues $\alpha$ and $\beta$. Show that any eigenvector corresponding to $\alpha$ is orthogonal to any eigenvector corresponding to $\beta$.
(*Nagoya University*)

## Hint.

Two vectors $\mathrm{u}$ and $\mathrm{v}$ are orthogonal if their inner (dot) product $\mathrm{u} \cdot \mathrm{v} := \mathrm{u}^{\mathrm{T}}\mathrm{v} = 0$.

Here $\mathrm{u}^{\mathrm{T}}$ is the transpose of $\mathrm{u}$.

A fact that we will use below is that for matrices $A$ and $B$, we have $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

### Proof.

Let $\mathrm{u}, \mathrm{v}$ be eigenvectors corresponding to $\alpha, \beta$, respectively.

Namely we have

$$A\mathrm{u} = \alpha\mathrm{u} \text{ and } A\mathrm{v} = \beta\mathrm{v}. \tag{*}$$

To prove that $\mathrm{u}$ and $\mathrm{v}$ are orthogonal, we show that the inner product $\mathrm{u} \cdot \mathrm{v} = 0$.

Keeping this in mind, we compute

$$\begin{aligned}
&\alpha(\mathrm{u} \cdot \mathrm{v}) = (\alpha\mathrm{u}) \cdot \mathrm{v} \\
&\overset{(*)}{=} A\mathrm{u} \cdot \mathrm{v} = (A\mathrm{u})^{\mathrm{T}}\mathrm{v} \\
&= \mathrm{u}^{\mathrm{T}}A^{\mathrm{T}}\mathrm{v} \text{ (This follows from the fact mentioned in the hint above)} \\
&= \mathrm{u}^{\mathrm{T}}A\mathrm{v} \text{ (since } A \text{ is symmetric.)} \\
&\overset{(*)}{=} \mathrm{u}^{\mathrm{T}}\beta\mathrm{v} = \beta(\mathrm{u}^{\mathrm{T}}\mathrm{v}) = \beta(\mathrm{u} \cdot \mathrm{v}).
\end{aligned}$$

---

Therefore we obtain

$$\alpha(\mathrm{u} \cdot \mathrm{v}) = \beta(\mathrm{u} \cdot \mathrm{v}),$$

and thus

$$(\alpha - \beta)(\mathrm{u} \cdot \mathrm{v}) = 0.$$

Since $\alpha$ and $\beta$ are distinct, $\alpha - \beta \neq 0$.

Hence we must have

$$\mathrm{u} \cdot \mathrm{v} = 0,$$

and the eigenvectors $\mathrm{u}, \mathrm{v}$ are orthogonal.

㊹ Show that eigenvalues of a Hermitian matrix $A$ are real numbers.

### Proof 1.

Let $\lambda$ be an arbitrary eigenvalue of a Hermitian matrix $A$ and let $\mathrm{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$.

Then we have

$$A\mathrm{x} = \lambda\mathrm{x}. \tag{*}$$

Multiplying by $\bar{\mathrm{x}}^{\mathrm{T}}$ from the left, we obtain

$$\begin{aligned}
\bar{\mathrm{x}}^{\mathrm{T}}(A\mathrm{x}) &= \bar{\mathrm{x}}^{\mathrm{T}}(\lambda\mathrm{x}) \\
&= \lambda\bar{\mathrm{x}}^{\mathrm{T}}\mathrm{x} \\
&= \lambda||\mathrm{x}||.
\end{aligned}$$

We also have

$$\bar{\mathrm{x}}^{\mathrm{T}}(A\mathrm{x})=(A\mathrm{x})^{\mathrm{T}}\bar{\mathrm{x}}=\mathrm{x}^{\mathrm{T}}A^{\mathrm{T}}\bar{\mathrm{x}}.$$

The first equality follows because the dot product $\mathrm{u}\cdot\mathrm{v}$ of vectors $\mathrm{u},\mathrm{v}$ is commutative.

That is, we have

$$\mathrm{u}\cdot\mathrm{v}=\mathrm{u}^{\mathrm{T}}\mathrm{v}=\mathrm{v}^{\mathrm{T}}\mathrm{u}=\mathrm{v}\cdot\mathrm{u}.$$

We applied this fact with $\mathrm{u}=\bar{\mathrm{x}}$ and $\mathrm{v}=A\mathrm{x}$ .

---

Thus we obtain

$$\mathrm{x}^{\mathrm{T}}A^{\mathrm{T}}\bar{\mathrm{x}}=\lambda||\mathrm{x}||.$$

Taking the complex conjugate of this equality, we have

$$\bar{\mathrm{x}}^{\mathrm{T}}\bar{A}^{\mathrm{T}}\mathrm{x}=\bar{\lambda}||\mathrm{x}||. \tag{**}$$

(Note that $\bar{\bar{\mathrm{x}}}=\mathrm{x}$. Also $\overline{||\mathrm{x}||}=||\mathrm{x}||$ because $||\mathrm{x}||$ is a real number.)

Since the matrix $A$ is Hermitian, we have $\bar{A}^{\mathrm{T}}=A$ .

---

This yields that

$$\begin{aligned}\bar{\lambda}||\mathrm{x}|| &\overset{(**)}{=} \bar{\mathrm{x}}^{\mathrm{T}}A\mathrm{x}\\ &\overset{(*)}{=} \bar{\mathrm{x}}^{\mathrm{T}}\lambda\mathrm{x}\\ &=\lambda||\mathrm{x}||.\end{aligned}$$

Recall that $\mathrm{x}$ is an eigenvector, hence $\mathrm{x}$ is not the zero vector and the length $||\mathrm{x}||\neq 0$.

Therefore, we divide by the length $||\mathrm{x}||$ and get

$$\lambda=\bar{\lambda}.$$

It follows from this that the eigenvalue $\lambda$ is a real number.
Since $\lambda$ is an arbitrary eigenvalue of $A$, we conclude that all the eigenvalues of the Hermitian matrix $A$ are real numbers.

## Proof 2.

Let $\lambda$ be an arbitrary eigenvalue of a Hermitian matrix $A$ and let $\mathrm{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$.
Then we have

$$A\mathrm{x}=\lambda\mathrm{x}. \tag{*}$$

Multiplying by $\bar{\mathrm{x}}^{\mathrm{T}}$ from the left, we obtain

$$\begin{aligned}\bar{\mathbf{x}}^{\mathrm{T}}(A\mathbf{x}) &= \bar{\mathbf{x}}^{\mathrm{T}}(\lambda\mathbf{x})\\ &= \lambda\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{x}\\ &= \lambda||\mathbf{x}||.\end{aligned}$$

Now we take the conjugate transpose of both sides and get

$$\bar{\mathbf{x}}^{\mathrm{T}}\bar{A}^{\mathrm{T}}\mathbf{x} = \bar{\lambda}||\mathbf{x}||. \tag{***}$$

Since $A$ is a Hermitian matrix, we have $\bar{A}^{\mathrm{T}} = A$.

Then the left hand side becomes

$$\begin{aligned}\bar{\mathbf{x}}^{\mathrm{T}}\bar{A}^{\mathrm{T}}\mathbf{x} &= \bar{\mathbf{x}}^{\mathrm{T}}A\mathbf{x}\\ &\overset{(*)}{=} \bar{\mathbf{x}}^{\mathrm{T}}\lambda\mathbf{x}\\ &= \lambda||\mathbf{x}||. \end{aligned} \tag{****}$$

Therefore comparing (***) and (****) we obtain

$$\lambda||\mathbf{x}|| = \bar{\lambda}||\mathbf{x}||.$$

Since $\mathbf{x}$ is an eigenvector, it is not the zero vector and the length $||\mathbf{x}|| \neq 0$.

Dividing by the length $||\mathbf{x}||$, we obtain $\lambda = \bar{\lambda}$ and this implies that $\lambda$ is a real number.
Since $\lambda$ is an arbitrary eigenvalue of $A$, we conclude that every eigenvalue of the Hermitian matrix $A$ is a real number.

## Corollary

Every real symmetric matrix is Hermitian. Thus, as a corollary of the problem we obtain the following fact:

Eigenvalues of a real symmetric matrix are real.

## Related Question.

**Problem**. Let $A$ be an $n \times n$ real symmetric matrix.
Prove that there exists an eigenvalue $\lambda$ of $A$ such that for any vector $\mathbf{v} \in \mathbb{R}^n$, we have the inequality

$$\mathbf{v} \cdot A\mathbf{v} \leq \lambda||\mathbf{v}||^2.$$

Note that the inequality makes sense because eigenvalues of $A$ are real by Corollary.

For a proof of this problem, see the post “Inequality about Eigenvalue of a Real Symmetric Matrix“.

㊺ Prove that the matrix

$$A = \begin{bmatrix} 1 & 1.00001 & 1 \\ 1.00001 & 1 & 1.00001 \\ 1 & 1.00001 & 1 \end{bmatrix}$$

has one positive eigenvalue and one negative eigenvalue.
(*University of California, Berkeley*)

### Solution.

Let us put $a = 1.00001$. We compute the characteristic polynomial $\det(A - tI)$ of the given matrix $A$ as follows.
We have

$$\begin{aligned} \det(A - tI) &= \begin{vmatrix} 1-t & a & 1 \\ a & 1-t & a \\ 1 & a & 1-t \end{vmatrix} \\ &= (1-t)\begin{vmatrix} 1-t & a \\ a & 1-t \end{vmatrix} - a\begin{vmatrix} a & a \\ 1 & 1-t \end{vmatrix} + \begin{vmatrix} a & 1-t \\ 1 & a \end{vmatrix} \end{aligned}$$

by the cofactor expansion corresponding to the first row.

Simplifying this, we obtain

$$\det(A - tI) = -t(t^2 - 3t + 2 - 2a^2).$$

---

The eigenvalues of $A$ are roots of this characteristic polynomial.
Hence $0$ is an eigenvalue of $A$. Let $\lambda_1$ and $\lambda_2$ be other two eigenvalues of $A$.

Then we have

$$\det(A - tI) = -t(t - \lambda_1)(t - \lambda_2) = -t(t^2 - (\lambda_1 + \lambda_2) + \lambda_1 \lambda_2).$$

Therefore we have

$$\lambda_1 \lambda_2 = 2 - 2a^2.$$

Since $2 - 2a^2 = 2(1 - a^2) < 0$ as $a = 1.0001 > 1$, the product $\lambda_1 \lambda_2$ is negative, and we conclude that one of them is positive and the other is negative.

(Note that if the constant $c$ term of a quadratic polynomial $x^2 + bx + c$ is negative, then the roots of the polynomial are real and one is negative and the other is positive.)

In summary, the eigenvalues of the matrix $A$ are $0$ and one positive eigenvalue and one negative eigenvalue.

㊻ Suppose that $\lambda$ and $\mu$ are two distinct eigenvalues of a square matrix $A$ and let $\mathbf{x}$ and $\mathbf{y}$ be eigenvectors corresponding to $\lambda$ and $\mu$, respectively. If $a$ and $b$ are nonzero numbers, then prove that $a\mathbf{x} + b\mathbf{y}$ is not an eigenvector of $A$ (corresponding to any eigenvalue of $A$

## Hint.

We use the following fact in the proof.

**Fact**: Two eigenvectors corresponding to distinct eigenvalues are linearly independent.

## Proof.

Seeking a contradiction, we assume that $a\mathrm{x} + b\mathrm{y}$ is an eigenvector corresponding to an eigenvalue $\zeta$. Thus we have

$$A(a\mathrm{x} + b\mathrm{y}) = \zeta(a\mathrm{x} + b\mathrm{y}). \tag{*}$$

---

We calculate the left hand side of this equality as follows.

We have

$$\begin{aligned} A(a\mathrm{x} + b\mathrm{y}) &= aA\mathrm{x} + bA\mathrm{y} \\ &= a\lambda\mathrm{x} + b\mu\mathrm{y} \end{aligned}$$

since $A\mathrm{x} = \lambda\mathrm{x}, A\mathrm{y} = \mu\mathrm{y}$ by defintion.

Therefore, from (*) we obtain

$$a(\lambda - \zeta)\mathrm{x} + b(\mu - \zeta)\mathrm{y} = 0.$$

---

Recall that eigenvectors corresponding to distinct eigenvalues are linearly independent. Thus $\mathrm{x}$ and $\mathrm{y}$ are linearly independent.

Thus, the coefficients of the above linear combinations must be zero:

$$a(\lambda - \zeta) = 0 \text{ and } b(\mu - \zeta) = 0.$$

Since $a \neq 0, b \neq 0$, this implies that we have

$$\lambda = \zeta = \mu,$$

and this is a contradiction because $\lambda$ and $\mu$ are supposed to be distinct.

Hence, $a\mathrm{x} + b\mathrm{y}$ cannot be an eigenvector of any eigenvalue of $A$ .

(47) Consider a polynomial

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_i$ are real numbers. Define the matrix

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{bmatrix}.$$

Then prove that the characteristic polynomial $\det(xI - A)$ of $A$ is the polynomial $p(x)$. The matrix is called the **companion matrix** of the polynomial $p(x)$.

---

㊽ Let $A$ be an $n \times n$ matrix. Suppose that all the eigenvalues $\lambda$ of $A$ are real and satisfy $\lambda < 1$. Then show that the determinant $\det(I - A) > 0$, where $I$ is the $n \times n$ identity matrix.

---

㊾ Consider the $2 \times 2$ matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

where $\theta$ is a real number $0 \leq \theta < 2\pi$.
**(a)** Find the characteristic polynomial of the matrix $A$.
**(b)** Find the eigenvalues of the matrix $A$.
**(c)** Determine the eigenvectors corresponding to each of the eigenvalues of $A$.

---

## Proof.

### (a) Find the characteristic polynomial of the matrix $A$.

The characteristic polynomial $p(t)$ of $A$ is computed as follows.

Let $I$ be the $2 \times 2$ identity matrix.

We have

$$\begin{aligned} p(t) &= \det(A - tI) \\ &= \begin{vmatrix} \cos\theta - t & -\sin\theta \\ \sin\theta & \cos\theta - t \end{vmatrix} \\ &= (\cos\theta - t)^2 + \sin^2\theta \qquad (*) \\ &= t^2 - (2\cos\theta)t + \cos^2\theta + \sin^2\theta \\ &= t^2 - (2\cos\theta)t + 1 \end{aligned}$$

by the trigonometry identity $\cos^2\theta + \sin^2\theta = 1$.

Hence the characteristic polynomial of $A$ is

$$p(t) = t^2 - (2\cos\theta)t + 1.$$

### (b) Find the eigenvalues of the matrix $A$.

The eigenvalues of $A$ are roots of the characteristic polynomial $p(t)$.

So let us solve

$$p(t) = t^2 - (2\cos\theta)t + 1 = 0.$$

By the quadratic formula, we have

$$\begin{aligned} t &= \frac{2\cos\theta \pm \sqrt{(2\cos\theta)^2 - 4}}{2} \\ &= \cos\theta \pm \sqrt{\cos^2\theta - 1} \\ &= \cos\theta \pm \sqrt{-\sin^2\theta} \\ &= \cos\theta \pm i\sin\theta = e^{\pm i\theta}. \end{aligned}$$

Hence eigenvalues of $A$ are

$$\cos\theta \pm i\sin\theta = e^{\pm i\theta}.$$

Alternatively, we could have used the equation (*).
Then we have

$$\begin{aligned} & p(t) = (\cos\theta - t)^2 + \sin^2\theta = 0 \\ & \Leftrightarrow (\cos\theta - t)^2 = -\sin^2\theta \\ & \Leftrightarrow \cos\theta - t = \pm i\sin\theta \end{aligned}$$

and thus, we obtain the eigenvalues

$$\cos\theta \pm i\sin\theta = e^{\pm i\theta}.$$

### (c) Determine the eigenvectors

Let us first find the eigenvectors corresponding to the eigenvalue $\lambda = \cos\theta + i\sin\theta$.
We have

$$A - \lambda I = \begin{bmatrix} -i\sin\theta & -\sin\theta \\ \sin\theta & -i\sin\theta \end{bmatrix}.$$

If $\theta = 0, \pi$, then $\sin\theta = 0$ and we have

$$A - \lambda I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and thus each nonzero vector of $\mathbb{R}^2$ is an eigenvector.

If $\lambda \neq 0, \pi$, then $\sin\theta \neq 0$.
Thus we have by elementary row operations

$$A - \lambda I = \begin{bmatrix} -i\sin\theta & -\sin\theta \\ \sin\theta & -i\sin\theta \end{bmatrix} \xrightarrow[\frac{1}{\sin\theta}R_2]{\frac{i}{\sin\theta}R_1} \begin{bmatrix} 1 & -i \\ 1 & -i \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}.$$

It follows that the eigenvectors for $\lambda$ are

$$\begin{bmatrix} i \\ 1 \end{bmatrix} t,$$

for any nonzero scalar $t$.

The other eigenvalue is the conjugate $\bar{\lambda}$ of $\lambda$.
Since $A$ is a real matrix, the eigenvectors of $\bar{\lambda}$ are the conjugate of those of $\lambda$.
Hence the eigenvectors corresponding to $\bar{\lambda}$ is

$$\begin{bmatrix} -i \\ 1 \end{bmatrix} t,$$

for any nonzero scalar $t$.

In summary, when $\theta = 0, \pi$, the eigenvalues are $1, -1$, respectively, and every nonzero vector of $\mathbb{R}^2$ is an eigenvector.

If $\theta \neq 0, \pi$, then the eigenvectors corresponding to the eigenvalue $\cos\theta + i\sin\theta$ are

$$\begin{bmatrix} i \\ 1 \end{bmatrix} t,$$

for any nonzero scalar $t$.
The eigenvectors corresponding to the eigenvalue $\cos\theta - i\sin\theta$ are

$$\begin{bmatrix} -i \\ 1 \end{bmatrix} t,$$

for any nonzero scalar $t$.

## Related Question.

For a similar proble, try the following.

**Problem.**
For a real number $0 \leq \theta \leq \pi$, we define the real $3 \times 3$ matrix $A$ by

$$A = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**(a)** Find the determinant of the matrix $A$.

**(b)** Show that $A$ is an orthogonal matrix.

**(c)** Find the eigenvalues of $A$.

㊿ Let $A$ be an $n \times n$ matrix and let $\lambda_1, \dots, \lambda_n$ be its eigenvalues. Show that
**(a)** $\det(A) = \prod_{i=1}^{n} \lambda_i$.
**(b)** $\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i$.

## Plan 1.

(51) Use the definition of eigenvalues (the characteristic polynomial).

(52) Compare coefficients.

### Plan 2.

(53) Make $A$ upper triangular matrix or in the Jordan normal/canonical form.

(54) Use the property of determinants and traces.

### Proof. [Method 1]

**(1)** Recall that eigenvalues are roots of the characteristic polynomial $p(\lambda) = \det(A - \lambda I_n)$ .
It follows that we have

$$\det(A - \lambda I_n) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1,n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{m2} & \cdots & a_{nn} - \lambda \end{vmatrix} = \prod_{i=1}^{n} (\lambda_i - \lambda). \qquad (*)$$

Letting $\lambda = 0$, we see that $\det(A) = \prod_{i=1}^{n} \lambda_i$ and this completes the proof of part (a).

**(2)** Compare the coefficients of $\lambda^{n-1}$ of the both sides of (*).
The coefficient of $\lambda^{n-1}$ of the determinant on the left side of (*) is

$$(-1)^{n-1}(a_{11} + a_{22} + \cdots a_{nn}) = (-1)^{n-1}\mathrm{tr}(A).$$

The coefficient of $\lambda^{n-1}$ of the determinant on the right side of (*) is

$$(-1)^{n-1}\sum_{i=1}^{n} \lambda_i .$$

Thus we have $\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$ .

### Proof. [Method 2]

Observe that there exists an $n \times n$ invertible matrix $P$ such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}. \qquad (**)$$

This is an upper triangular matrix and diagonal entries are eigenvalues.
(If this is not familiar to you, then study a "triangularizable matrix" or "Jordan normal/canonical form".)

**(1)** Since the determinant of an upper triangular matrix is the product of diagonal entries, we have

$$\begin{aligned}\prod_{i=1}^{n} \lambda_i &= \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) \\ &= \det(P)^{-1}\det(A)\det(P) = \det(A),\end{aligned}$$

where we used the multiplicative property of the determinant.

---

**(2)** We take the trace of both sides of (**) and get

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr}(P^{-1}AP) = \operatorname{tr}(A).$$

(Here for the last equality we used the property of the trace that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for any $n \times n$ matrices $A$ and $B$.)
Thus we obtained the result $\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i$.

## Comment.

The proof of (1) in the first method is simple, but that of (2) requires a bit observation, especially when we find the coefficient of the left-hand side.

The proof of (2) in the second method is simpler although you need to know about the Jordan normal/canonical form.

These two formulas relate the determinant and the trace, and the eigenvalue of a matrix in a very simple way.

55 **(a)** A $2 \times 2$ matrix $A$ satisfies $\operatorname{tr}(A^2) = 5$ and $\operatorname{tr}(A) = 3$.
Find $\det(A)$.
**(b)** A $2 \times 2$ matrix has two parallel columns and $\operatorname{tr}(A) = 5$. Find $\operatorname{tr}(A^2)$.
**(c)** A $2 \times 2$ matrix $A$ has $\det(A) = 5$ and positive integer eigenvalues. What is the trace of $A$?
(*Harvard University*)

## Solution.

For (a) and (b), we give two solutions. The first one does not use the knowledge of eigenvalues, and the second one uses eigenvalues.

## Solution 1 of (a) A $2 \times 2$ matrix $A$ satisfies $\operatorname{tr}(A^2) = 5$ and $\operatorname{tr}(A) = 3$. Find $\det(A)$.

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then we have

$$A^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{bmatrix}.$$

Since $\operatorname{tr}(A^2) = 5$ and $\operatorname{tr}(A) = 3$, we obtain

$$\begin{aligned} &5 = \operatorname{tr}(A^2) = (a^2 + bc) + (bc + d^2) = a^2 + 2bc + d^2 \text{ and} \\ &3 = \operatorname{tr}(A) = a + d. \end{aligned}$$

We find the determinant $\det(A) = ad - bc$ as follows.

We have

$$\begin{aligned} \det(A) = ad - bc &= \frac{1}{2}\big((a+d)^2 - (a^2 + 2bc + d^2)\big) \\ &= \frac{1}{2}(3^2 - 5) = 2. \end{aligned}$$

Thus, we obtain $\det(A) = 2$.

## Solution 2 of (a)

Let $\lambda_1$ and $\lambda_2$ be eigenvalues of $A$.

Then we have

$$\begin{aligned} 3 = \operatorname{tr}(A) = \lambda_1 + \lambda_2 \text{ and} \\ 5 = \operatorname{tr}(A^2) = \lambda_1^2 + \lambda_2^2. \end{aligned}$$

Here we used two facts.

The first one is that the trace of a matrix is the sum of all eigenvalues of the matrix.

The second one is that $\lambda^2$ is an eigenvalue of $A^2$ if $\lambda$ is an eigenvalue of $A$, and these are all the eigenvalues of $A^2$.

Since the determinant of $A$ is the product of eigenvalues of $A$, we have

$$\begin{aligned} \det(A) &= \lambda_1 \lambda_2 \\ &= \frac{1}{2}\big((\lambda_1 + \lambda_2)^2 - (\lambda_1^2 + \lambda_2^2)\big) \\ &= \frac{1}{2}(3^2 - 5) \\ &= 2. \end{aligned}$$

Hence we have $\det(A) = 2$.

Solution 1 of (b) A $2 \times 2$ matrix has two parallel columns and $\mathrm{tr}(A) = 5$. Find $\mathrm{tr}(A^2)$.

Since two columns are parallel, we can write $A$ as

$$A = \begin{bmatrix} a & ra \\ c & rc \end{bmatrix}.$$

Then we have

$$5 = \mathrm{tr}(A) = a + rc.$$

We use the formula in Solution 1 of (a) for $\mathrm{tr}(A^2)$ with $b = ra$ and $d = rc$, and we compute

$$\begin{aligned}\mathrm{tr}(A^2) &= a^2 + 2(ra) + (rc)^2 \\ &= (a + rc)^2 \\ &= 5^2 = 25.\end{aligned}$$

Thus, we find $\mathrm{tr}(A^2) = 25$.

Solution 2 of (b)

Since two columns are parallel, the matrix $A$ is singular. Hence $A$ has an eigenvalue $0$.

Since the sum of all the eigenvalues is $\mathrm{tr}(A) = 5$, we see that $0$ and $5$ are eigenvalues of $A$.

It follows that $0$ and $25$ are eigenvalues of $A^2$. Hence

$$\mathrm{tr}(A^2) = 0 + 25 = 25.$$

Solution of (c) A $2 \times 2$ matrix $A$ has $\det(A) = 5$ and positive integer eigenvalues. What is the trace of $A$?

The product of eigenvalues of $A$ is the determinant $\det(A) = 5$.

Since eigenvalues are positive integers, it follows that $1$ and $5$ are eigenvalues of $A$.

It follows that

$$\mathrm{tr}(A) = 1 + 5 = 6.$$

(56) Let $n$ be an odd integer and let $A$ be an $n \times n$ real matrix. Prove that the matrix $A$ has at least one real eigenvalue.

## Proof 1.

Let $p(t) = \det(A - tI)$ be the characteristic polynomial of the matrix $A$.

It is a degree $n$ polynomial and the coefficients are real numbers since $A$ is a real matrix. Since $n$ is odd, the leading term of $p(t)$ is $-t^n$. That is, we have

$$p(t) = -t^n + \text{lower terms}.$$

(Note: if you use the alternative definition of characteristic polynomial $p(t) = \det(tI - A)$ , then the leading term is $t^n$ . You can easily modify the proof for this case.)

Therefore, as $t$ increases the polynomial $p(t)$ tends to $-\infty$:

$$\lim_{t\to\infty} = -\infty.$$

Similarly, we have

$$\lim_{t\to-\infty} = \infty.$$

By the intermediate value theorem, there is $-\infty < \lambda < \infty$ such that$p(\lambda) = 0$ . Since a root of $p(t)$ is an eigenvalue of $A$ , we have obtained a real eigenvalue $\lambda$ of $A$ .

## Proof 2.

Let $p(t) = \det(A - tI)$ be the characteristic polynomial of the matrix $A$ and write it as

$$p(t) = \prod_{i=1}^{k} (\lambda_i - t)^{n_i},$$

where $\lambda_i$ are distinct eigenvalues of $A$ and $n_i$ is the algebraic multiplicity of $\lambda_i$ .

Since $A$ is a real matrix, the coefficients of $p(t)$ are real numbers. So we have

$$\overline{p(t)} = p(\bar{t}).$$

Thus, we have

$$p(\bar{t}) = \overline{p(t)} = \prod_{i=1}^{k} (\bar{\lambda}_i - \bar{t})^{n_i}.$$

Replacing $\bar{t}$ by $t$, we obtain

$$p(t) = \prod_{i=1}^{k} (\bar{\lambda}_i - t)^{n_i}.$$

This yields that the complex conjugate $\bar{\lambda}_i$ is an eigenvalue with algebraic multiplicity $n_i$ .
Hence each non-real eigenvalue $\lambda$ appears as a pair $(\lambda, \bar{\lambda})$ counting multiplicity.
Thus, there are even number of non-real eigenvalues of $A$ .

The number of all eigenvalues of $A$ counting multiplicities is $n$ .
Since $n$ is odd, there must be at least one real eigenvalue of $A$ .

57 Let

$$A = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-3} & a_{n-2} \\ a_{n-2} & a_{n-1} & \cdots & a_{n-4} & a_{n-3} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_2 & a_3 & \cdots & a_0 & a_1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{bmatrix}$$

be a complex $n \times n$ matrix. Such a matrix is called **circulant** matrix. Prove that the determinant of the circulant matrix $A$ is given by

$$\det(A) = \prod_{k=0}^{n-1} (a_0 + a_1\zeta^k + a_2\zeta^{2k} + \cdots + a_{n-1}\zeta^{k(n-1)}),$$

where $\zeta = e^{2\pi i/n}$ is a primitive $n$-th root of unity.

## Proof.

Let $\omega$ be any $n$-th root of unity.

Consider the vector

$$\mathrm{v} = \begin{bmatrix} 1 \\ \omega \\ \omega^2 \\ \vdots \\ \omega^{n-1} \end{bmatrix}.$$

We show that the vector $\mathrm{v}$ is an eigenvector of $A$.

We compute

$$A\mathrm{v} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-3} & a_{n-2} \\ a_{n-2} & a_{n-1} & \cdots & a_{n-4} & a_{n-3} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_2 & a_3 & \cdots & a_0 & a_1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{bmatrix} \begin{bmatrix} 1 \\ \omega \\ \omega^2 \\ \vdots \\ \omega^{n-1} \end{bmatrix}.$$

The first entry of the vector $A\mathrm{v}$ is

$$a_0 + a_1\omega + a_2\omega^2 + \cdots a_{n-2}\omega^{n-2} + a_{n-1}\omega^{n-1} =: \lambda.$$

We define $\lambda$ to be this number.

---

The second entry is

$$\begin{aligned} & a_{n-1} + a_0\omega + \cdots + a_{n-3}\omega^{n-2} + a_{n-2}\omega^{n-1} \\ & = (a_{n-1}\omega^{n-1} + a_0 + \cdots + a_{n-3}\omega^{n-3} + a_{n-2}\omega^{n-2})\omega \\ & = (a_0 + \cdots + a_{n-3}\omega^{n-3} + a_{n-2}\omega^{n-2} + a_{n-1}\omega^{n-1})\omega \\ & = \lambda\omega \end{aligned}$$

---

Similarly the $i$-th entry of the vector $A\mathrm{v}$ is

$$\begin{aligned}&a_{n-i+1} + a_{n-i+2}\,\omega + \cdots + a_{n-i}\,\omega^{n-1}\\ &= (a_{n-i+1}\,\omega^{n-i+1} + a_{n-i+2}\,\omega^{n-i+2} + \cdots + a_{n-i}\,\omega^{n-i})\omega^{i-1}\\ &= \lambda\omega^{i-1}.\end{aligned}$$

Therefore we obtain

$$A\mathbf{v} = \begin{bmatrix} \lambda \\ \lambda\omega \\ \lambda\omega^2 \\ \vdots \\ \lambda\omega^{n-1} \end{bmatrix} = \lambda\mathbf{v}.$$

---

Since $\mathbf{v}$ is a nonzero vector, it follows that $\lambda$ is an eigenvalue of $A$ and $\mathbf{v}$ is an eigenvector corresponding to $\lambda$.

---

The above argument holds for any $n$-th root of unity $\omega$.
We take $\omega = \zeta^k$, where $k$ runs from $0$ to $n-1$.
It follows that the vector

$$\mathbf{v}_k := \begin{bmatrix} 1 \\ \zeta^k \\ \zeta^{2k} \\ \vdots \\ \zeta^{k(n-1)} \end{bmatrix}$$

is eigenvector corresponding to the eigenvalue

$$\lambda_k := a_0 + a_1\zeta^k + a_2\zeta^{2k} + \cdots a_{n-2}\zeta^{k(n-2)} + a_{n-1}\zeta^{k(n-1)}$$

for each $k = 0, 1, \dots, n-1$.

---

We claim that the vectors $\mathbf{v}_k$ are linearly independent.

To see this, form a matrix whose column vectors are these vectors. That is, we consider

$$B = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \zeta & \zeta^2 & \dots & \zeta^{n-2} & \zeta^{n-1} \\ 1 & \zeta^2 & \zeta^4 & \dots & \zeta^{2(n-2)} & \zeta^{2(n-1)} \\ 1 & \zeta^3 & \zeta^6 & \dots & \zeta^{3(n-2)} & \zeta^{3(n-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 1 & \zeta^{n-1} & \zeta^{2(n-1)} & \dots & \zeta^{(n-1)(n-2)} & \zeta^{(n-1)(n-1)} \end{bmatrix}.$$

---

This is the Vandermonde matrix and its determinant is

$$\det(B) = \prod_{i<j} (\zeta^j - \zeta^i) \neq 0.$$

Thus, the matrix $B$ is nonsingular, hence its column vectors are linearly independent. It follows that $\lambda_k$, $k = 0, 1, \dots, n$ are all the eigenvalues of $A$. Since the determinant is the product of all the eigenvalues of $A$, we have

$$\begin{aligned}\det(A) &= \prod_{k=0}^{n-1} \lambda_k \\ &= \prod_{k=0}^{n-1} (a_0 + a_1\zeta^k + a_2\zeta^{2k} + \cdots + a_{n-1}\zeta^{k(n-1)}),\end{aligned}$$

as required. This completes the proof.

(58) Let $A$ be an $n \times n$ real matrix. Prove the followings:
**(a)** The matrix $AA^{\mathrm{T}}$ is a symmetric matrix.
**(b)** The set of eigenvalues of $A$ and the set of eigenvalues of $A^{\mathrm{T}}$ are equal.
**(c)** The matrix $AA^{\mathrm{T}}$ is non-negative definite.
(An $n \times n$ matrix $B$ is called *non-negative definite* if for any $n$ dimensional vector $\mathbf{x}$, we have $\mathbf{x}^{\mathrm{T}} B\mathbf{x} \geq 0$.)
**(d)** All the eigenvalues of $AA^{\mathrm{T}}$ is non-negative.

## Facts.

Before the proofs, we first review several basic properties of the transpose of a matrix.

(59) For any matrices $A$ and $B$ so that the product $AB$ is defined, we have $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$

(60) We have $(A^{\mathrm{T}})^{\mathrm{T}} = A$ for any matrix $A$.

Also recall that the eigenvalues of a matrix $A$ are the solutions of the characteristic polynomial $p_A(x) = \det(A - xI)$ .

## Proof.

### (a) The matrix $AA^{\mathrm{T}}$ is a symmetric matrix.

We compute $(AA^{\mathrm{T}})^{\mathrm{T}} = (A^{\mathrm{T}})^{\mathrm{T}} A^{\mathrm{T}} = AA^{\mathrm{T}}$ and thus $AA^{\mathrm{T}}$ is a symmetric matrix.

(We used the Fact 1 and 2.)

### (b) The set of eigenvalues of $A$ and the set of eigenvalues of $A^{\mathrm{T}}$ are equal.

We show that the characteristic polynomials of $A$ and $A^{\mathrm{T}}$ are the same, hence they have exactly same eigenvalues.
Let $p_A(x)$ and $p_{A^{\mathrm{T}}}(x)$ be the characteristic polynomials of $A$ and $A^{\mathrm{T}}$, respectively. Then we have

$$p_A(x) = \det(A - xI) = \det(A - xI)^{\mathrm{T}} = \det(A^{\mathrm{T}} - xI) = p_{A^{\mathrm{T}}}(x).$$

The first and last equalities are the definition of the characteristic polynomial.
The second equality is true because in general we have $\det(B) = \det(B^{\mathrm{T}})$ for a square matrix $B$. This completes the proof of (b).

### (c) The matrix $AA^{\mathrm{T}}$ is non-negative definite.

Let $\mathbf{x}$ be an $n$ dimensional vector. Then we have

$$\mathbf{x}^{\mathrm{T}}AA^{\mathrm{T}}\mathbf{x} = (A^{\mathrm{T}}\mathbf{x})^{\mathrm{T}}(A^{\mathrm{T}}\mathbf{x}) = ||A^{\mathrm{T}}\mathbf{x}|| \geq 0,$$

since a norm (length) of a vector is always non-negative.
Thus $AA^{\mathrm{T}}$ is non-negative definite.

### (d) All the eigenvalues of $AA^{\mathrm{T}}$ is non-negative.

Let $\lambda$ be an eigenvalue of $AA^{\mathrm{T}}$ and let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$.
Then we compute

$$\mathbf{x}^{\mathrm{T}}AA^{\mathrm{T}}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\lambda\mathbf{x} = \lambda||\mathbf{x}||.$$

Here the first equality follows from the definitions of the eigenvalue $\lambda$ and eigenvector $\mathbf{x}$. In part (c), we proved that $AA^{\mathrm{T}}$ is non-negative definite, hence we have $\lambda||\mathbf{x}|| \geq 0$.
Therefore $\lambda \geq 0$ and this completes the proof of (d).

## Related Question.

Problem. A real symmetric $n \times n$ matrix $A$ is called **positive definite** if

$$\mathbf{x}^{\mathrm{T}}A\mathbf{x} > 0$$

for all nonzero vectors $\mathbf{x}$ in $\mathbb{R}^n$.

**(a)** Prove that the eigenvalues of a real symmetric positive-definite matrix $A$ are all positive.

**(b)** Prove that if eigenvalues of a real symmetric matrix $A$ are all positive, then $A$ is positive-definite.

(61) Let $A$ be an $n \times n$ matrix. Suppose that $\mathbf{y}$ is a nonzero row vector such that $\mathbf{y}A = \mathbf{y}$. (Here a row vector means a $1 \times n$ matrix.) Prove that there is a nonzero column vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{x}$. (Here a column vector means an $n \times 1$ matrix.)

## Proof 1.(Without the theory of eigenvalues)

Let $I$ be the $n \times n$ identity matrix. Then we have

$$0_{1\times n} = \mathbf{y}A - \mathbf{y} = \mathbf{y}(A - I),$$

where $0_{1\times n}$ is the row zero vector.

Taking the transpose, we have

$$\begin{aligned} 0_{n\times 1} &= 0_{1\times n}^{\mathrm{T}} = (\mathbf{y}(A-I))^{\mathrm{T}} \\ &= (A-I)^{\mathrm{T}}\mathbf{y}^{\mathrm{T}}. \end{aligned}$$

Since the vector $\mathbf{y}$ is nonzero, the transpose $\mathbf{y}^{\mathrm{T}}$ is a nonzero column vector.
Thus, the above equality yields that the matrix $(A-I)^{\mathrm{T}}$ is singular.
It follows that the matrix $A-I$ is singular as well.
Hence there exists a nonzero column vector $\mathbf{x}$ such that

$$(A-I)\mathbf{x} = 0_{n\times 1},$$

and consequently we have

$$A\mathbf{x} = \mathbf{x}$$

for a nonzero column vector $\mathbf{x}$.

## Proof 2. (Using the theory of eigenvalues)

Taking the conjugate of the both sides of the identity $\mathbf{y}A = \mathbf{y}$, we obtain

$$A^{\mathrm{T}}\mathbf{y}^{\mathrm{T}} = \mathbf{y}^{\mathrm{T}}.$$

Since $\mathbf{y}$ is a nonzero row vector, $\mathbf{y}^{\mathrm{T}}$ is a nonzero column vector.
It follows that $1$ is an eigenvalue of the matrix $A^{\mathrm{T}}$ and $\mathbf{y}^{\mathrm{T}}$ is a corresponding eigenvector.

Since the matrices $A$ and $A^{\mathrm{T}}$ has the same eigenvalues , we deduce that the matrix $A$ has an eigenvalue $1$.
(See part (b) of the post "Transpose of a matrix and eigenvalues and related questions.".)
Let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $1$ (by definition $\mathbf{x}$ is nonzero). Then we have

$$A\mathbf{x} = \mathbf{x}$$

as required.

62 **(a)** Let $A$ be a real orthogonal $n \times n$ matrix. Prove that the length (magnitude) of each eigenvalue of $A$ is $1$.
**(b)** Let $A$ be a real orthogonal $3 \times 3$ matrix and suppose that the determinant of $A$ is $1$. Then prove that $A$ has $1$ as an eigenvalue.

## Proof.

### (a) Prove that the length (magnitude) of each eigenvalue of $A$ is $1$

Let $A$ be a real orthogonal $n \times n$ matrix.
Let $\lambda$ be an eigenvalue of $A$ and let $\mathbf{v}$ be a corresponding eigenvector.
Then we have

$$A\mathrm{v} = \lambda \mathrm{v}.$$

It follows from this we have

$$\|A\mathrm{v}\|^2 = \|\lambda \mathrm{v}\|^2 = |\lambda|^2 \|\mathrm{v}\|^2.$$

The left hand side becomes

$$\begin{aligned}
&\|A\mathrm{v}\|^2 \\
&= \overline{(A\mathrm{v})}^{\mathrm{T}}(A\mathrm{v}) && \text{by definition of the length} \\
&= \bar{\mathrm{v}}^{\mathrm{T}} A^{\mathrm{T}} A\mathrm{v} && \text{because } A \text{ is real} \\
&= \bar{\mathrm{v}}^{\mathrm{T}} \mathrm{v} && \text{because } A^{\mathrm{T}}A = I \text{ as } A \text{ is orthogonal} \\
&= \|\mathrm{v}\|^2 && \text{by definition of the length.}
\end{aligned}$$

It follows that we obtain

$$\|\mathrm{v}\|^2 = |\lambda|^2 \|\mathrm{v}\|^2.$$

Since $\mathrm{v}$ is an eigenvector, it is non-zero, and hence $\|\mathrm{v}\| \neq 0$.
Canceling $\|\mathrm{v}\|$, we have

$$|\lambda|^2 = 1.$$

Since the length is non-negative, we obtain

$$|\lambda| = 1,$$

as required.

> (b) Prove that $A$ has $1$ as an eigenvalue.

Let $A$ be a real orthogonal $3 \times 3$ matrix with $\det(A) = 1$.
Let us consider the characteristic polynomial $p(t) = \det(A - tI)$ of $A$.
The roots of $p(t)$ are eigenvalues of $A$.

Since $A$ is a real $3 \times 3$ matrix, the degree of the polynomial $p(t)$ is $3$ and the coefficients are real.
Thus, there are two cases to consider:

(63) there are three real eigenvalues $\alpha, \beta, \gamma$, and

(64) there is one real eigenvalue $\alpha$ and a complex conjugate pair $\beta, \bar{\beta}$ of eigenvalues.

Let us first deal with case 1.
By part (a), the lengths of eigenvalues $\alpha, \beta, \gamma$ are $1$. Since they are real numbers, we have

$$\alpha = \pm 1, \beta = \pm 1, \gamma = \pm 1.$$

Recall that the product of all eigenvalues of $A$ is the determinant of $A$.
(For a proof, see the post "Determinant/trace and eigenvalues of a matrix".)
Thus we have

$$\alpha\beta\gamma = \det(A) = 1.$$

Thus, at least one of $\alpha, \beta, \gamma$ is $1$.

Next, we consider case 2. Again the lengths of eigenvalues $\alpha, \beta, \bar{\beta}$ are $1$.
Then we have

$$\begin{aligned} 1 &= \det(A) = \alpha\beta\bar{\beta} \\ &= \alpha|\beta|^2 = \alpha. \end{aligned}$$

Therefore, in either case, we see that $A$ has $1$ as an eigenvalue.

(65) Let $A$ and $B$ be square matrices such that they commute each other: $AB = BA$.
Assume that $A - B$ is a nilpotent matrix. Then prove that the eigenvalues of $A$ and $B$ are the same.

## Proof.

Let $N := A - B$ . By assumption, the matrix $N$ is nilpotent.
This means that there exists a positive integer $n$ such that $N^n$ is the zero matrix $O$.

Let $\lambda$ be an eigenvalue of $B$ and let $\mathrm{v}$ be an eigenvector corresponding to $\lambda$. That is, we have $B\mathrm{v} = \lambda\mathrm{v}$ and $\mathrm{v} \neq 0$.
We prove that $\lambda$ is also an eigenvalue of $A$.

---

Note that since $A$ and $B$ commute each other, it follows that the matrices $N$ and $B - \lambda I$ commute each other as well.
Then we compute

$$\begin{aligned} (A - \lambda I)^n &= (N + B - \lambda I)^n \\ &= \sum_{i=0}^{n} \binom{n}{i} N^i (B - \lambda I)^{n-i}, \end{aligned}$$

where the second equality follows by the binomial expansion. (Note that the binomial expansion is true for matrices commuting each other.)
Then we have

$$(A - \lambda I)^n \mathrm{v} = \sum_{i=0}^{n-1} \binom{n}{i} N^i (B - \lambda I)^{n-i} \mathrm{v} + N^n \mathrm{v} = 0$$

since $(B - \lambda I)\mathrm{v} = 0$ and $N^n = O$.

---

This implies that there exists an integer $k$, $0 \leq k \leq n-1$ such that

$$\mathrm{u} := (A - \lambda I)^k \mathrm{v} \neq 0 \text{ and } (A - \lambda I)^{k+1} \mathrm{v} = 0.$$

It yields that $(A-\lambda I)\mathbf{u}=0$ and $\mathbf{u}\neq 0$, or equivalently $A\mathbf{u}=\lambda\mathbf{u}$ .
Hence $\lambda$ is an eigenvalue of $A$ .

This proves that each eigenvalue of $B$ is an eigenvalue of $A$ .
Note that if $A-B$ is nilpotent, then $B-A$ is also nilpotent.
Thus, switching the roles of $A$ and $B$ , we also see that each eigenvalue of $A$ is an eigenvalue of $B$ .
Therefore, the eigenvalues of $A$ and $B$ are the same.

(66) Let $A$ be an $n\times n$ matrix. Suppose that $A$ has real eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ with corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$. Furthermore, suppose that $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Let

$$\mathbf{x}_0 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$$

for some real numbers $c_1, c_2, \dots, c_n$ and $c_1 \neq 0$. Define $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for k=0, 1, 2,\dots and let

$$\beta_k = \frac{\mathbf{x}_k\cdot\mathbf{x}_{k+1}}{\mathbf{x}_k\cdot\mathbf{x}_k} = \frac{\mathbf{x}_k^{\mathrm{T}}\mathbf{x}_{k+1}}{\mathbf{x}_k^{\mathrm{T}}\mathbf{x}_k}.$$

Prove that $\lim_{k\to\infty}\beta_k = \lambda_1$.

### Proof.

We have

$$x_k = A\mathbf{x}_{k-1} = A^2\mathbf{x}_{k-2} = \cdots = A^k\mathbf{x}_0$$

by applying the definition of $\mathbf{x}_k$ successively.
Then we have

$$\begin{aligned}\mathbf{x}_k &= A^k\mathbf{x}_0\\ &= A^k(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n)\\ &= c_1A^k\mathbf{u}_1 + c_2A^k\mathbf{u}_2 + \cdots + c_nA^k\mathbf{u}_n\\ &= c_1\lambda_1^k\mathbf{u}_1 + c_2\lambda_2^k\mathbf{u}_2 + \cdots + c_n\lambda_n^k\mathbf{u}_n,\end{aligned}$$

where the last step follows from the equalities $A^k\mathbf{u}_i = \lambda_i^k\mathbf{u}_i$, which in turn follows from the definitions of eigenvalues and eigenvectors $A\mathbf{u}_i = \lambda_i\mathbf{u}_i$ .
Using the sum notation, we simply write it as

$$\mathbf{x}_k = \sum_{i=1}^{n} c_i\lambda_i^k\mathbf{u}_i.$$

This formula is true for any nonnegative integer $k$.

Let us next compute the dot product $\mathbf{x}_k\cdot\mathbf{x}_k$. We have

$$
\begin{aligned}
\mathbf{x}_k \cdot \mathbf{x}_k &= \left(\sum_{i=1}^{n} c_i \lambda_i^k \mathbf{u}_i\right) \cdot \left(\sum_{j=1}^{n} c_j \lambda_j^k \mathbf{u}_j\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \lambda_i^k \lambda_j^k \mathbf{u}_i \cdot \mathbf{u}_j \\
&= \lambda_1^{2k} \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \left(\frac{\lambda_i}{\lambda_1}\right)^k \left(\frac{\lambda_j}{\lambda_1}\right)^k \mathbf{u}_i \cdot \mathbf{u}_j .
\end{aligned}
$$

Since $\lambda_1$ is strictly larger than any other eigenvalues, we have

$$
\lim_{k\to\infty} \left(\frac{\lambda_i}{\lambda_1}\right)^k = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{if } i = 2, 3, \ldots, n. \end{cases}
$$

It follows that

$$
\lim_{k\to\infty} \mathbf{x}_k \cdot \mathbf{x}_k = \lambda_1^{2k} c_1^2 \mathbf{u}_1 \cdot \mathbf{u}_1 .
$$

---

Similarly, we compute

$$
\begin{aligned}
\mathbf{x}_k \cdot \mathbf{x}_{k+1} &= \left(\sum_{i=1}^{n} c_i \lambda_i^k \mathbf{u}_i\right) \cdot \left(\sum_{j=1}^{n} c_j \lambda_j^{k+1} \mathbf{u}_j\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \lambda_i^k \lambda_j^{k+1} \mathbf{u}_i \cdot \mathbf{u}_j \\
&= \lambda_1^{2k+1} \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \left(\frac{\lambda_i}{\lambda_1}\right)^k \left(\frac{\lambda_j}{\lambda_1}\right)^{k+1} \mathbf{u}_i \cdot \mathbf{u}_j .
\end{aligned}
$$

Thus we obtain

$$
\lim_{k\to\infty} \mathbf{x}_k \cdot \mathbf{x}_{k+1} = \lambda_1^{2k+1} c_1^2 \mathbf{u}_1 \cdot \mathbf{u}_1 .
$$

---

Combining these computations, we have

$$
\begin{aligned}
\lim_{k\to\infty} \beta_k &= \frac{\lim_{k\to\infty} \mathbf{x}_k \cdot \mathbf{x}_{k+1}}{\lim_{k\to\infty} \mathbf{x}_k \cdot \mathbf{x}_k} \\
&= \frac{\lambda_1^{2k+1} c_1^2 \mathbf{u}_1 \cdot \mathbf{u}_1}{\lambda_1^{2k} c_1^2 \mathbf{u}_1 \cdot \mathbf{u}_1} \\
&= \lambda_1
\end{aligned}
$$

as required. This completes the proof.

# Eigenvectors and Eigenspaces

**Definition**

Let $A$ be an $n \times n$ matrix.

1. The eigenspace corresponding to an eigenvalue $\lambda$ of $A$ is defined to be $E_\lambda = \{\mathbf{x} \in \mathbb{C}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$.

**Summary**

Let $A$ be an $n \times n$ matrix.

1. The eigenspace $E_\lambda$ consists of all eigenvectors corresponding to $\lambda$ and the zero vector.
2. $A$ is singular if and only if $0$ is an eigenvalue of $A$.
3. The nullity of $A$ is the geometric multiplicity of $\lambda = 0$ if $\lambda = 0$ is an eigenvalue.

## Problems

1. Let $A$ be a $3 \times 3$ matrix. Suppose that $A$ has eigenvalues $2$ and $-1$, and suppose that $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors corresponding to $2$ and $-1$, respectively, where

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Then compute $A^5\mathbf{w}$, where $\mathbf{w} = \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix}$.

**Solution.**

Since $\mathbf{u}$ is an eigenvector corresponding to the eigenvalue $2$, we have

$$A\mathbf{u} = 2\mathbf{u}.$$

Similarly, we have

$$A\mathbf{v} = -\mathbf{v}.$$

From these, we have

$$A^5\mathbf{u} = 2^5\mathbf{u} \text{ and } A\mathbf{v} = (-1)^5\mathbf{v}.$$

To compute $A^5\mathbf{w}$, we first need to express $\mathbf{w}$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$. Thus, we need to find scalars $c_1, c_2$ such that

$$\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}.$$

By inspection, we have

$$\begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix},$$

and thus we obtain $c_1 = 3$ and $c_2 = 2$,

We compute $A^5\mathbf{w}$ as follows:

$$\begin{aligned} A^5\mathbf{w} &= A^5(3\mathbf{u} + 2\mathbf{v}) \\ &= 3A^5\mathbf{u} + 2A^5\mathbf{v} \\ &= 3 \cdot 2^5\mathbf{u} + 2 \cdot (-1)^5\mathbf{v} \\ &= 96\mathbf{u} - 2\mathbf{v} \\ &= 96\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 92 \\ -2 \\ -96 \end{bmatrix}. \end{aligned}$$

Therefore, the result is

$$A^5\mathbf{w} = \begin{bmatrix} 92 \\ -2 \\ -96 \end{bmatrix}.$$

---

**2** Let $A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 4 & 1 \\ 2 & -4 & 0 \end{bmatrix}$. The matrix $A$ has an eigenvalue $2$. Find a basis of the eigenspace $E_2$ corresponding to the eigenvalue $2$.
(*The Ohio State University*)

---

### Solution.

By definition, the eigenspace $E_2$ corresponding to the eigenvalue $2$ is the null space of the matrix $A - 2I$

.

That is, we have

$$E_2 = \mathcal{N}(A - 2I).$$

We reduce the matrix $A - 2I$ by elementary row operations as follows.

$$A-2I=\begin{bmatrix}-1 & 2 & 1\\ -1 & 2 & 1\\ 2 & -4 & -2\end{bmatrix}$$

$$\xrightarrow[R_3+2R_1]{R_2-R_1}\begin{bmatrix}-1 & 2 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\xrightarrow{-R_1}\begin{bmatrix}1 & -2 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

Thus, the solutions $\mathbf{x}$ of $(A-2I)\mathbf{x}=0$ satisfy $x_1=2x_2+x_3$.

Thus, the null space $\mathcal{N}(A-2I)$ consists of vectors

$$\mathbf{x}=\begin{bmatrix}2x_2+x_3\\ x_2\\ x_3\end{bmatrix}=x_2\begin{bmatrix}2\\ 1\\ 0\end{bmatrix}+x_3\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$$

for any scalars $x_2, x_3$.

---

Hence we have

$$E_2=\mathcal{N}(A-2I)=\operatorname{Span}\left(\begin{bmatrix}2\\ 1\\ 0\end{bmatrix},\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}\right).$$

It is straightforward to see that the vectors $\begin{bmatrix}2\\ 1\\ 0\end{bmatrix},\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$ are linearly independent, hence they form a basis of $E_2$.

Thus, a basis of $E_2$ is

$$\left\{\begin{bmatrix}2\\ 1\\ 0\end{bmatrix},\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}\right\}.$$

③ Let

$$A=\begin{bmatrix}0 & 0 & 0 & 0\\ 1 & 1 & 1 & 1\\ 0 & 0 & 0 & 0\\ 1 & 1 & 1 & 1\end{bmatrix}.$$

One of the eigenvalues of the matrix $A$ is $\lambda=0$. Find the geometric multiplicity of the eigenvalue $\lambda=0$.

## Solution (a). Eigenvalues of $A$ and algebraic multiplies

Eigenvalues and their algebraic multiplicities are determined by the characteristic polynomial $p(t)$ of $A$.

By definition, the characteristic polynomial of $A$ is $p(t)=\det(A-tI)$ .

We have

$$\begin{aligned}
p(t) &= \det(A - tI) \\
&= \begin{vmatrix} -t & 0 & 0 & 0 \\ 1 & 1-t & 1 & 1 \\ 0 & 0 & -t & 0 \\ 1 & 1 & 1 & 1-t \end{vmatrix} \\
&= -t \begin{vmatrix} 1-t & 1 & 1 \\ 0 & -t & 0 \\ 1 & 1 & 1-t \end{vmatrix} && \text{by the first row cofactor expansion} \\
&= -t\left( -t \begin{vmatrix} 1-t & 1 \\ 1 & 1-t \end{vmatrix} \right) && \text{by the second row cofactor expansion} \\
&= t^2\left( (1-t)^2 - 1 \right) \\
&= t^2(t^2 - 2t) \\
&= t^3(t-2).
\end{aligned}$$

Thus the characteristic polynomial is

$$p(t) = t^3(t-2).$$

From this, the eigenvalues of $A$ are $0$ and $2$ with algebraic multiplicities $3$ and $1$, respectively.

## Solution (b). Geometric multiplicity

We give two solutions for part (b).

### First Solution (b). (Finding the rank first)

Recall that the geometric multiplicity of $\lambda$ is the dimension of the eigenspace $E_\lambda = \mathcal{N}(A - \lambda I)$. That is, the geometric multiplicity of $\lambda$ is the nullity of the matrix $A - \lambda I$.

Let us now consider the case $\lambda = 0$.

We first find the rank of $A - 0I = A$ as follows.

$$A - 0I = A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_4 - R_2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The last matrix is in reduced row echelon form.

Hence the rank of $A$ is $1$.

The rank-nullity theorem says that

$$\text{rank of } A + \text{nullity of } A = 4.$$

Thus, the nullity of $A = A - 0I$ is $3$, and hence the geometric multiplicity of $\lambda = 0$ is $3$.

### Second Solution (b). (Finding a basis of the eigenspace)

In this solution, we find a basis of the eigenspace $E_0$.

By definition $E_0 = \mathcal{N}(A - 0I) = \mathcal{N}(A)$.

Thus, the eigenspace $E_0$ is the null space of the matrix $A$.

We solve the equation $A\mathrm{x} = 0$ as follows.

The augmented matrix of this equation is

$$[A \mid 0] = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_4 - R_2} \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence the solution satisfies

$$x_1 = -x_2 - x_3 - x_4$$

and the general solution is

$$\mathrm{x} = \begin{bmatrix} -x_2 - x_3 - x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, the eigenspace is

$$\begin{aligned} E_0 &= \mathcal{N}(A) \\ &= \left\{ \mathrm{x} \in \mathbb{C}^4 \;\middle|\; \mathrm{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \text{ for any } x_2, x_3, x_4 \in \mathbb{C} \right\} \\ &= \operatorname{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Thus the set

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a spanning set of $E_0$, and it is straightforward to check that the set is linearly independent.

Hence this set is a basis of $E_0$, and the dimension of $E_0$ is $3$.

The geometric multiplicity of $\lambda = 0$ is the dimension of $E_0$ by definition.

Thus, the geometric multiplicity of $\lambda$ is $3$.

## Comment.

These are Quiz 12 problems for Math 2568 (Introduction to Linear Algebra) at OSU in Spring 2017.

I could have combined these two problems into one problem, and just asked to find eigenvalues and algebraic/ geometric multiplicities for each eigenvalue.

The reason I didn't do is that I wanted to rescue students who didn't get a correct answer in part (a) for some reasons.

Also, since one of the eigenvalue is given in (b), students could use this information to double check their solutions in (a).

(At least if you didn't get the eigenvalue $0$, you made a mistake somewhere.)

④ Let $A$ and $B$ be $n \times n$ matrices. Suppose that these matrices have a common eigenvector $\mathbf{x}$. Show that $\det(AB - BA) = 0$.

## Steps.

⑤ Write down eigenequations of $A$ and $B$ with the eigenvector $\mathrm{x}$.

⑥ Show that AB-BA is singular.

⑦ A matrix is singular if and only if the determinant of the matrix is zero.

## Proof.

Let $\alpha$ and $\beta$ be eigenvalues of $A$ and $B$ such that the vector $\mathrm{x}$ is a corresponding eigenvector.

Namely we have $A\mathrm{x} = \alpha\mathrm{x}$ and $B\mathrm{x} = \beta\mathrm{x}$ .

Then we have

$$\begin{aligned}(AB - BA)\mathrm{x} &= AB\mathrm{x} - BA\mathrm{x} = A(\beta\mathrm{x}) - B(\alpha\mathrm{x}) \\ &= \beta A\mathrm{x} - \alpha B\mathrm{x} = \beta\alpha - \alpha\beta = 0.\end{aligned}$$

By the definition of eigenvector, $\mathrm{x}$ is a non-zero vector. Thus the matrix $AB - BA$ is singular.

Equivalently the determinant of $AB - BA$ is zero.

## Comment.

This is a simple necessary condition that $A$ and $B$ have a common eigenvector.

Here are few derived questions.

⑧ Is this a sufficient condition?

⑨ If so prove it.

⑩ If not give a counterexample,

11 and find a necessary and sufficient condition.

12 Suppose that $A$ is a diagonalizable matrix with characteristic polynomial

$$f_A(\lambda) = \lambda^2(\lambda-3)(\lambda+2)^3(\lambda-4)^3.$$

**(a)** Find the size of the matrix $A$.
**(b)** Find the dimension of $E_4$, the eigenspace corresponding to the eigenvalue $\lambda = 4$.
**(c)** Find the dimension of the nullspace of $A$.
(*Stanford University*)

## Hint.

If a matrix is diagonalizable, then the algebraic multiplicity of an eigenvalue is the same as the geometric multiplicity of the eigenvalue.

Recall that the geometric multiplicity of an eigenvalue is the dimension of the eigenspace of the eigenvalue.

## Solution.

### (a) Find the size of the matrix $A$

If $A$ is an $n \times n$ matrix, then its characteristic polynomial is of degree $n$. Since the degree of $f_A$ is $9$, the size of $A$ is $9 \times 9$.

### (b) Find the dimension of $E_4$, the eigenspace corresponding to the eigenvalue $\lambda = 4$.

The algebraic multiplicity of the eigenvalue $\lambda = 4$ is $3$. (This is the number of times that the factor $\lambda - 4$ appears in $f_A$.)
Since $A$ is diagonalizable, the algebraic multiplicity is the same as the geometric multiplicity for each eigenvalue.
The geometric multiplicity is the dimension of the eigenspace by definition. Thus the dimension of $E_4$ is $3$.

### (c) Find the dimension of the kernel(nullspace) of $A$.

Note that $\ker(A) = \{x \in \mathbb{R}^9 \mid Ax = 0\} = \{x \in \mathbb{R}^9 \mid (A - 0I_9)x = 0\} = E_0$. Thus the kernel of $A$ is the same as the eigenspace $E_0$ corresponding to the eigenvalue $\lambda = 0$.
Since $A$ is diagonalizable, the geometric multiplicity of $\lambda = 0$, which is the dimension of the eigenspace $E_0 = \ker(A)$, is the same as the algebraic multiplicity of $\lambda = 0$.
Thus the dimension of $\ker(A)$ is $2$.

13 Let $A$ be a square matrix and its characteristic polynomial is give by

$$p(t) = (t-1)^3(t-2)^2(t-3)^4(t-4).$$

Find the rank of $A$.
(*The Ohio State University*)

## Solution.

Note that the degree of the characteristic polynomial $p(t)$ is the size of the matrix $A$.
Since the degree of $p(t)$ is $3+2+4+1=10$, the size of the matrix $A$ is $10\times 10$.

From the characteristic polynomial, we see that the eigenvalues of $A$ are $1,2,3,4$.
In particular, $0$ is not an eigenvalue of $A$.
Hence the null space of $A$ is zero dimensional, that is, the nullity of $A$ is $0$.

By the rank-nullity theorem, we have

$$\text{rank of } A + \text{ nullity of } A = 10.$$

Hence the rank of $A$ is $10$.

14 **(a)** Let

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

be a matrix such that $a_{11}+a_{12}=1$ and $a_{21}+a_{22}=1$. Namely, the sum of the entries in each row is $1$. (Such a matrix is called (right) ***stochastic matrix***.) Then prove that the matrix $A$ has an eigenvalue $1$.
**(b)** Find all the eigenvalues of the matrix

$$B=\begin{bmatrix} 0.3 & 0.7 \\ 0.6 & 0.4 \end{bmatrix}.$$

**(c)** For each eigenvalue of $B$, find the corresponding eigenvectors.

## Hint.

15 For (a), consider the vector $\mathbf{x}=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

16 For (b), use (a) and consider the trace of $B$ and its relation to eigenvalues.
For this relation, see the problem Determinant/trace and eigenvalues of a matrix.

## Solution.

### (a) Prove that the matrix $A$ has an eigenvalue $1$.

Let $\mathbf{x}=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and we compute

$$A\mathbf{x}=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}=\begin{bmatrix} a_{11}+a_{12} \\ a_{21}+a_{22} \end{bmatrix}=\begin{bmatrix} 1 \\ 1 \end{bmatrix}=1\cdot\mathbf{x}.$$

This shows that $A$ has the eigenvalue $1$.

### (b) Find all the eigenvalues of the matrix $B$.

Note that the matrix $B$ is of the type of the matrix in (a).
Thus the matrix $B$ has the eigenvalue $1$. Since $B$ is $2$ by $2$ matrix, it has two eigenvalues counting multiplicities. To find the other eigenvalue, we note that the trace is the sum of the eigenvalues.

Thus we have

$$\operatorname{tr}(B) = 0.3 + 0.4 = 1 + \lambda,$$

where $\lambda$ is the second eigenvalue. Hence another eigenvalue is $\lambda = -0.3$.

### (c) For each eigenvalue of $B$, find the corresponding eigenvectors.

By solving $(B - I)\mathrm{x} = 0$ and $(B + 0.3I)\mathrm{x} = 0$, we find that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} t \quad \text{and} \quad \begin{bmatrix} -7 \\ 6 \end{bmatrix} t$$

for any nonzero scalar $t$ are eigenvectors corresponding to eigenvalues $1$ and $-0.3$, respectively.

### Comment.

For some specific matrices, we can find eigenvalues without solving the characteristic polynomials like we did in part (b).

17 Let $A$ be an $n \times n$ matrix. Suppose that $\lambda_1, \lambda_2$ are distinct eigenvalues of the matrix $A$ and let $\mathbf{v}_1, \mathbf{v}_2$ be eigenvectors corresponding to $\lambda_1, \lambda_2$, respectively. Show that the vectors $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent.

### Proof.

To show that the vectors $\mathrm{v}_1, \mathrm{v}_2$ are linearly independent, consider a linear combination

$$c_1\mathrm{v}_1 + c_2\mathrm{v}_2 = 0, \tag{*}$$

where $c_1, c_2$ are complex numbers.
Our goal is to show that $c_1 = c_2 = 0$.

By the definitions of eigenvalues and eigenvectors we have

$$A\mathrm{v}_1 = \lambda_1\mathrm{v}_1 \text{ and } A\mathrm{v}_2 = \lambda_2\mathrm{v}_2.$$

Multiplying $A$ and (*), we have

$$\begin{aligned} 0 &= A \cdot 0 \\ &= A(c_1\mathrm{v}_1 + c_2\mathrm{v}_2) \\ &= c_1A\mathrm{v}_1 + c_2A\mathrm{v}_2 \\ &= c_1\lambda_1\mathrm{v}_1 + c_2\lambda_2\mathrm{v}_2. \end{aligned} \tag{**}$$

Now we multiply $\lambda_2$ and (*) and we obtain

$$0 = c_1\lambda_2\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2.$$

We subtract (**) from the last expression, and get

$$\begin{aligned}0 &= (c_1\lambda_2\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2) - (c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2)\\ &= c_1(\lambda_2 - \lambda_1)\mathbf{v}_1.\end{aligned}$$

Recall that an eigenvector is by definition a nonzero vector, and hence $\mathbf{v}_1 \neq \mathbf{0}$.
Thus we must have

$$c_1(\lambda_2 - \lambda_1) = 0.$$

Since $\lambda_1$ and $\lambda_2$ are distinct, we must have $c_1 = 0$.

Substituting $c_1 = 0$ into (*), we also see that $c_2 = 0$ since $\mathbf{v}_2 \neq \mathbf{0}$.
Therefore, the values of $c_1$ and $c_2$ are both zero, and hence the eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent.

## Related Question.

As an application of this problem, try the next problem, which provides a geometrical meaning of eigenvalues.

**Problem**. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation and let $A$ be the matrix representation of $T$ with respect to the standard basis of $\mathbb{R}^2$.

Prove that the following two statements are equivalent.

⑱ There are exactly two distinct lines $L_1, L_2$ in $\mathbb{R}^2$ passing through the origin that are mapped onto themselves:

$$T(L_1) = L_1 \text{ and } T(L_2) = L_2.$$

⑲ The matrix $A$ has two distinct nonzero real eigenvalues.

For a proof of this problem, see the post “A Linear Transformation Preserves Exactly Two Lines If and Only If There are Two Real Non-Zero Eigenvalues“.

⑳ Let $A$ and $B$ be an $n \times n$ matrices. Suppose that all the eigenvalues of $A$ are distinct and the matrices $A$ and $B$ commute, that is $AB = BA$. Then prove that each eigenvector of $A$ is an eigenvector of $B$.

## Hint.

Each eigenspace for $A$ is one dimensional.

Proof.

Since $A$ has $n$ distinct eigenvalues, the characteristic polynomial for $A$ factors into the product of degree $1$ polynomials.

Thus, the algebraic multiplicity of each eigenvalue is, $1$ and hence the geometric multiplicity is also $1$. (The geometric multiplicity is always less than or equal to the algebraic multiplicity and greater than 0 by definition.)

Thus the dimension of each eigenspace, which is the geometric multiplicity, is $1$.

---

Let $\lambda$ be an eigenvalue of the matrix $A$ and let $\mathbf{x}$ be the eigenvector corresponding to $\lambda$.
Since the eigenspace $E_\lambda$ for $\lambda$ is one dimensional and $\mathbf{x} \in E_\lambda$ is a nonzero vector in it, the vector $\mathbf{x}$ is a basis.
That is, we have $E_\lambda = \{t\mathbf{x} \mid t \in \mathbb{C}\}$ .

---

Now we multiply $A\mathbf{x} = \lambda\mathbf{x}$ by the matrix $B$ on the left and obtain

$$\begin{aligned} & BA\mathbf{x} = \lambda B\mathbf{x} \\ \iff \quad & AB\mathbf{x} = \lambda B\mathbf{x} \text{ since } AB = BA. \end{aligned}$$

---

This implies that $B\mathbf{x} \in E_\lambda = \{t\mathbf{x} \mid t \in \mathbb{C}\}$ . Therefore there exists $t \in \mathbb{C}$ such that $B\mathbf{x} = t\mathbf{x}$ .

Hence the vector $\mathbf{x}$ is also an eigenvector corresponding to the eigenvalue $t$ of the matrix $B$.
This completes the proof.

㉑ Let $H$ and $E$ be $n \times n$ matrices satisfying the relation $HE - EH = 2E$. Let $\lambda$ be an eigenvalue of the matrix $H$ such that the real part of $\lambda$ is the largest among the eigenvalues of $H$. Let $\mathbf{x}$ be an eigenvector corresponding to $\lambda$. Then prove that $E\mathbf{x} = \mathbf{0}$.

Proof.

Using the given relation we have

$$\begin{aligned} HE\mathbf{x} - EH\mathbf{x} &= 2E\mathbf{x} \\ \Rightarrow HE\mathbf{x} - \lambda E\mathbf{x} &= 2E\mathbf{x} \\ \Rightarrow HE\mathbf{x} &= (\lambda + 2)E\mathbf{x}. \end{aligned} \tag{*}$$

Let $\mathbf{v} = E\mathbf{x}$ and assume that $\mathbf{v} \neq \mathbf{0}$. Then we have from (*)

$$H\mathbf{v} = (\lambda + 2)\mathbf{v} \text{ for the nonzero vector } \mathbf{v}.$$

This means that $\lambda + 2$ is an eigenvalue of the matrix $H$ and $\mathbf{v}$ is a corresponding eigenvector.

However, the real part of the eigenvalue $\lambda + 2$ is greater than that of $\lambda$. This contradicts the choice of $\lambda$.
Therefore the vector $\mathbf{v} = E\mathbf{x}$ must be zero.

Comment.

You might wonder why we consider the relation $HE - EH = 2E$.
In fact, this relation is a part of the relations of the Lie algebra $\mathrm{sl}(2)$.

Although you don't have to know anything about Lie algebra to solve this problem,
it might be good to know that these computations are actually used in a more advanced mathematic.

---

**22** Let

$$A = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}.$$

Pick your favorite number $a$. Find the dimension of the null space of the matrix $A - aI$, where $I$ is the $3 \times 3$ identity matrix. Your score of this problem is equal to that dimension times five.
(*The Ohio State University*)

---

## Solution.

Note that if $a$ is not an eigenvalue of the matrix $A$, then the matrix $A - aI$ is nonsingular, that is, the null space of $A - aI$ is the zero space, and the dimension is $0$.
Thus to obtain a nonzero point, your favorite number should be an eigenvalue.

Let us find the eigenvalue. The characteristic polynomial $p(t)$ of $A$ is

$$\begin{aligned} p(t) &= \det(A - tI) \\ &\begin{vmatrix} 5-t & 2 & -1 \\ 2 & 2-t & 2 \\ -1 & 2 & 5-t \end{vmatrix} \\ &= (5-t)\begin{vmatrix} 2-t & 2 \\ 2 & 5-t \end{vmatrix} - 2\begin{vmatrix} 2 & 2 \\ -1 & 5-t \end{vmatrix} + (-1)\begin{vmatrix} 2 & 2-t \\ -1 & 2 \end{vmatrix} \\ &\text{(By the first row cofactor expansion.)} \\ &= (5-t)(t^2 - 7t + 6) - 2(-2t + 12) - (-t + 6) \\ &= -t^3 + 12t^2 - 36t \\ &= -t(t-6)^2. \end{aligned}$$

Eigenvalues are roots of the characteristic polynomial.
Thus eigenvalues of $A$ are $0$ and $6$, with algebraic multiplicities $1$ and $2$, respectively.

Since the matrix $A$ is a real symmetric matrix, it is diagonalizable. Thus the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity.

Therefore the dimension of the null space of $A - 6I$, which is the geometric multiplicity of the eigenvalue $6$ is equal to $2$, and the dimension of the null space of $A$ is $1$.

Thus, if you take $a = 6$, you get $2 \times 5 = 10$ points. If your favorite number is $a = 0$, then your score is $1 \times 5 = 5$ points. The other choice of $a$ will be zero points.

## Comment.

If you didn't notice that the given matrix is a real symmetric matrix, then you need to find eigenvectors corresponding to eigenvalues $6$ and to show that the eigenspace has dimension $2$.

㉓ Let $A = \begin{bmatrix} 1 & -14 & 4 \\ -1 & 6 & -2 \\ -2 & 24 & -7 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 4 \\ -1 \\ -7 \end{bmatrix}$. Find $A^{10}\mathbf{v}$. You may use the following information without proving it. The eigenvalues of $A$ are $-1, 0, 1$. The eigenspaces are given by

$$E_{-1} = \text{Span}\left\{ \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} \right\}, \quad E_0 = \text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} \right\}, \quad E_1 = \text{Span}\left\{ \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix} \right\}.$$

(*The Ohio State University*)

**Solution.**

Since $A$ has three distinct eigenvalues, the eigenvectors

$$\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}$$

form a basis of $\mathrm{R}^3$.

Thus, we can express the vector $\mathbf{v}$ as a linear combination

$$\mathbf{v} = x\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} + y\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} + z\begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}$$

for some scalars $x, y, z$.

We determine the values of $x, y, z$ by Gauss-Jordan elimination.

The augmented matrix of the system is

$$\begin{bmatrix} 3 & -2 & -4 & 4 \\ -1 & 1 & 2 & -1 \\ -5 & 4 & 7 & -7 \end{bmatrix}.$$

Applying the elementary row operations, we obtain

$$\begin{bmatrix} 3 & -2 & -4 & 4 \\ -1 & 1 & 2 & -1 \\ -5 & 4 & 7 & -7 \end{bmatrix} \underset{\text{Then } -R_1}{\xrightarrow{R_1 \leftrightarrow R_2}} \begin{bmatrix} 1 & -1 & -2 & 1 \\ 3 & -2 & -4 & 4 \\ -5 & 4 & 7 & -7 \end{bmatrix}$$

$$\underset{R_3+5R_1}{\xrightarrow{R_2-3R_1}} \begin{bmatrix} 1 & -1 & -2 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & -1 & -3 & -2 \end{bmatrix} \underset{\text{Then } -R_3}{\xrightarrow{\substack{R_1+R_2 \\ R_3+R_2}}} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow{R_2-2R_3} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

So we have

$$x = 2, y = -1, z = 1,$$

and the linear combination is

$$\mathbf{v} = 2\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}.$$

Recall that if $\lambda$ is an eigenvalue and $\mathbf{u}$ is a corresponding eigenvector of $A$, then we have

$$A^{10}\mathbf{u} = \lambda^{10}\mathbf{u}.$$

Using this property, we compute

$$A^{10}\mathbf{v} = A^{10}\left(2\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}\right)$$

$$= 2A^{10}\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} - A^{10}\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} + A^{10}\begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}$$

$$= 2(-1)^{10}\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} - 0^{10}\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} + (1)^{10}\begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}$$

$$= 2\begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix}$$

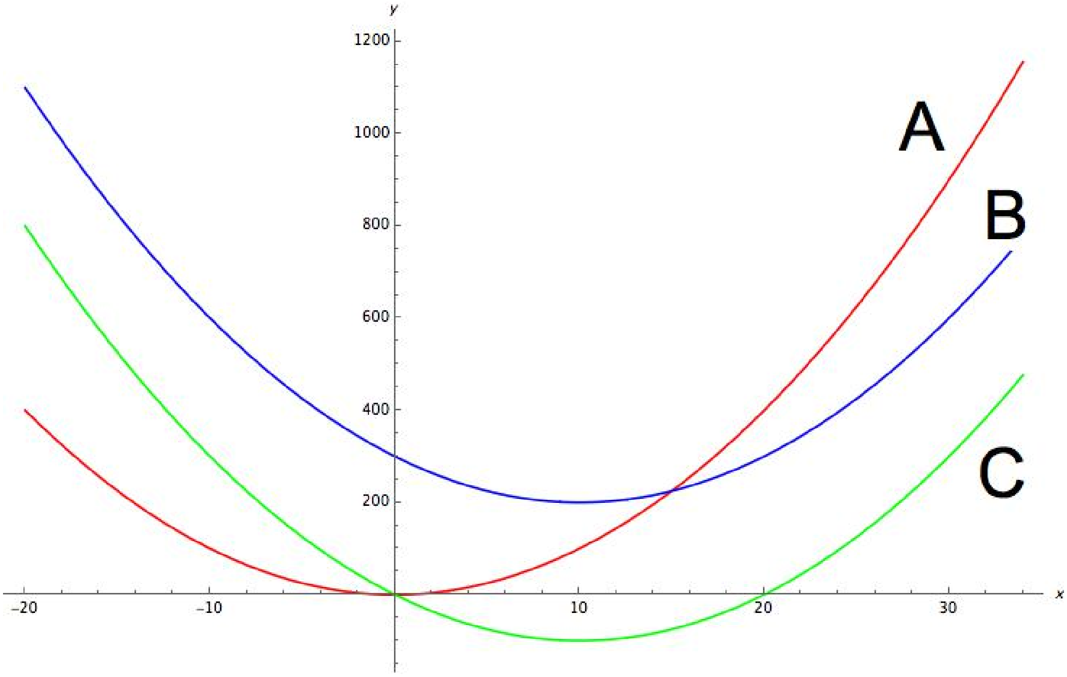
$$= \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix}.$$

In summary, we obtain

$$A^{10}\mathbf{v} = \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix}.$$

---

(24) Let $A, B, C$ are $2 \times 2$ diagonalizable matrices. The graphs of characteristic polynomials of $A, B, C$ are shown below. The red graph is for $A$, the blue one for $B$, and the green one for $C$. From this information, determine the rank of the matrices $A$, $B$, and $C$.



---

## Hint.

Observe that a null space a matrix $M$ is the same as the eigenspace $E_0$ corresponding to the eigenvalue $\lambda = 0$ (if 0 is an eigenvalue).
So what you need to see is whether the graphs pass though the origin.

Also, since the matrices are diagonalizable, the algebraic multiplicities are the same as the geometric multiplicities (the dimension of the eigenspace).
You can determine the algebraic multiplicities from the graph by looking at whether the graphs are tangential to $x$-axis or not.

## Solution.

We first determine the nullities of the matrices, and then using the rank-nullity theorem we obtain the ranks.

In general the nullity of any matrix $M$ is the dimension of the null space

$$\mathcal{N}(M) = \{\mathrm{x} \in \mathbb{R}^2 \mid M\mathrm{x} = 0\}$$

and the null space is the same as the eigenspace

$$E_0 = \{\mathrm{x} \in \mathbb{R}^2 \mid M\mathrm{x} = 0\mathrm{x} = 0\}$$

corresponding to the eigenvalue $\lambda = 0$ (if any).

From this observation, we see that the nullity of a matrix $M$ is the geometric multiplicity of the eigenspace $E_0$ associated to the eigenvalue $0$ (if any).

### The Nullity of the Matrix $A$

Now we determine the nullity of the matrix $A$.
The graph of the characteristic polynomial $p_A(\lambda)$ of $A$ passes through the origin $(0, 0)$.

Thus $\lambda = 0$ is a root of $p_A(\lambda)$ and hence $\lambda = 0$ is an eigenvalue.
Since the $x$-axis is tangential to the graph of $p_A(\lambda)$, the algebraic multiplicity of $\lambda = 0$ is $2$. Since the matrix $A$ is diagonalizable, the geometric multiplicity is the same as the algebraic multiplicity. Therefore the nullity of $A$ is $2$.

### The Nullity of the Matrix $B$

Next, the graph of characteristic polynomial $p_B(\lambda)$ of $B$ does not pass through the origin $(0, 0)$. Thus $\lambda = 0$ is not an eigenvalue of $B$. This yields that the matrix $B = B - 0I$ is nonsingular matrix, and hence we have the null space $\mathcal{N}(B) = \{0\}$ and the nullity of $B$ is zero.

### The Nullity of the Matrix $C$

Third, the origin $(0, 0)$ is on the graphs of characteristic polynomial $p_C(\lambda)$ of $C$, but the $x$-axis is not tangential to the graph of $p_C(\lambda)$.
Therefore the algebraic (hence geometric) multiplicity of $\lambda = 0$ is $1$. Thus the nullity of $C$ is $1$.

### Ranks by the Rank-Nullity Theorem

Finally, using the rank-nullity theorem

$$\text{rank} + \text{nullity} = 2,$$

we obtain the rank of $A, B, C$ are $0, 2, 1$, respectively.

㉕ Let $C$ be a $4 \times 4$ matrix with all eigenvalues $\lambda = 2, -1$ and eigensapces

$$E_2 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\} \text{ and } E_{-1} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

Calculate $C^4\mathbf{u}$ for $\mathbf{u} = \begin{bmatrix} 6 \\ 8 \\ 6 \\ 9 \end{bmatrix}$ if possible. Explain why if it is not possible!

(*The Ohio State University*)

## Solution.

Let

$$\mathrm{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathrm{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} \mathrm{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}.$$

The vector $\mathrm{v}_1$ is an eigenvector corresponding to eigenvalue $\lambda = 2$ and the vectors $\mathrm{v}_2, \mathrm{v}_2$ are eigenvectors corresponding to $\lambda = -1$.

---

If the vector $\mathrm{u}$ is a linear combination of $\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3$, then we can compute $C\mathrm{u}$ and hence $C^4\mathrm{u}$. So let us first determine whether

$$x_1\mathrm{v}_1 + x_2\mathrm{v}_2 + x_3\mathrm{v}_3 = \mathrm{u}$$

has a solution or not.

We use Gauss-Jordan elimination method to find the solution.
The augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & 2 & 1 & 8 \\ 1 & 1 & 1 & 6 \\ 1 & 1 & 2 & 9 \end{array}\right].$$

By elementary row operations, this matrix reduces to

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus the solution is

$$x_1 = 1, x_2 = 2, x_3 = 3$$

and we have the linear combination

$$\mathrm{v}_1 + 2\mathrm{v}_2 + 3\mathrm{v}_3 = \mathrm{u}.$$

---

Now we have

$$\begin{aligned} C^4\mathrm{u} &= C^4\mathrm{v}_1 + 2C^4\mathrm{v}_2 + 3C^4\mathrm{v}_3 \\ &= 2^4\mathrm{v}_1 + 2(-1)^4\mathrm{v}_2 + 3(-1)^4\mathrm{v}_3 \\ &= 16\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 21 \\ 23 \\ 21 \\ 24 \end{bmatrix}. \end{aligned}$$

Therefore we obtain

$$C^4 = \begin{bmatrix} 21 \\ 23 \\ 21 \\ 24 \end{bmatrix}.$$

26 Let $A$ be an $n \times n$ matrix and let $c$ be a complex number.
**(a)** For each eigenvalue $\lambda$ of $A$, prove that $\lambda + c$ is an eigenvalue of the matrix $A + cI$, where $I$ is the identity matrix. What can you say about the eigenvectors corresponding to $\lambda + c$?
**(b)** Prove that the algebraic multiplicity of the eigenvalue $\lambda$ of $A$ is the same as the algebraic multiplicity of the eigenvalue $\lambda + c$ of $A + cI$ are equal.
**(c)** How about geometric multiplicities?

## Proof.

### (a) $\lambda + c$ is an eigenvalue of $A + cI$.

Let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$. Then we have

$$A\mathbf{x} = \lambda\mathbf{x}.$$

It follows that we have

$$\begin{aligned} (A + cI)\mathbf{x} &= A\mathbf{x} + c\mathbf{x} \\ &= \lambda\mathbf{x} + c\mathbf{x} \\ &= (\lambda + c)\mathbf{x}. \end{aligned}$$

Thus, we obtain

$$(A + cI)\mathbf{x} = (\lambda + c)\mathbf{x},$$

where $\mathbf{x}$ is a nonzero vector.

Hence $\lambda + c$ is an eigenvalue of the matrix $A + cI$, and $\mathbf{x}$ is an eigenvector corresponding to $\lambda - c$.

In summary, if $\lambda$ is an eigenvalue of $A$ and $\mathbf{x}$ is an associated eigenvector, then $\lambda + c$ is an eigenvalue of $A + cI$ and $\mathbf{x}$ is an associated eigenvector corresponding to $\lambda + c$.

### (b) Algebraic multiplicities are the same

Let $p(t) = \det(A - tI)$ be the characteristic polynomial of $A$.
Let $q(t)$ be the characteristic polynomial of the matrix $A + cI$.
Then we have

$$\begin{aligned} q(t) &= \det\big((A + cI) - t\big) = \det\big(A - (t - c)\big) \\ &= p(t - c). \end{aligned}$$

Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of $A$ with algebraic multiplicities $n_1, \dots, n_k$, respectively.
Then we have

$$p(t) = \pm \prod_{i=1}^{k} (t - \lambda_i)^{n_i}.$$

It follows that we have

$$\begin{aligned} q(t) &= p(t-c) \\ &= \pm \prod_{i=1}^{k} (t-c-\lambda_i)^{n_i} \\ &= \pm \prod_{i=1}^{k} \left(t-(\lambda_i+c)\right)^{n_i}. \end{aligned}$$

From the last equation, we read that the eigenvalues of the matrix $A+cI$ are $\lambda_i + c$ with algebraic multiplicity $n_i$ for $i = 1, \dots, k$.
Thus, geometric multiplicities of $\lambda$ and $\lambda - c$ are the same.

### (c) How about geometric multiplicities?

From part (a), we know that eigenvectors of $\lambda$ are eigenvectors of $\lambda - c$.
Reversing the argument, the eigenvectors of $\lambda + c$ are eigenvectors of $\lambda$.

Thus, the eigenvectors correspond one to one, and the eigenspace of $\lambda$ is the same as the eigenspace of $\lambda + c$.
Hence their geometric multiplicities are the same.

## Applications

As applications of this problem, consider the following problems.

*Problem*. Find all the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 9 & 9 & 9 \\ 9 & 3 & 9 & 9 \\ 9 & 9 & 3 & 9 \\ 9 & 9 & 9 & 3 \end{bmatrix}.$$

This is a problem of a Linear Algebra final exam at Harvard University.
For a solution of this problem, see the post "Eigenvalues and eigenvectors of matrix whose diagonal entries are 3 and 9 elsewhere".

*Problem*. Find the determinant of the following matrix

$$A = \begin{bmatrix} 6 & 2 & 2 & 2 & 2 \\ 2 & 6 & 2 & 2 & 2 \\ 2 & 2 & 6 & 2 & 2 \\ 2 & 2 & 2 & 6 & 2 \\ 2 & 2 & 2 & 2 & 6 \end{bmatrix}.$$

This is also a problem of a Linear Algebra final exam at Harvard University.

See the post “Determinant of matrix whose diagonal entries are 6 and 2 elsewhere” for a solution.

---

㉗ Find all the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 9 & 9 & 9 \\ 9 & 3 & 9 & 9 \\ 9 & 9 & 3 & 9 \\ 9 & 9 & 9 & 3 \end{bmatrix}.$$

(*Harvard University*)

---

## Proof.

### (a) $\lambda + c$ is an eigenvalue of $A + cI$ .

Let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$. Then we have

$$A\mathbf{x} = \lambda\mathbf{x}.$$

It follows that we have

$$\begin{aligned}(A + cI)\mathbf{x} &= A\mathbf{x} + c\mathbf{x} \\ &= \lambda\mathbf{x} + c\mathbf{x} \\ &= (\lambda + c)\mathbf{x}.\end{aligned}$$

Thus, we obtain

$$(A + cI)\mathbf{x} = (\lambda + c)\mathbf{x},$$

where $\mathbf{x}$ is a nonzero vector.

Hence $\lambda + c$ is an eigenvalue of the matrix $A + cI$ , and $\mathbf{x}$ is an eigenvector corresponding to $\lambda - c$.

In summary, if $\lambda$ is an eigenvalue of $A$ and $\mathbf{x}$ is an associated eigenvector, then $\lambda + c$ is an eigenvalue of $A + cI$ and $\mathbf{x}$ is an associated eigenvector corresponding to $\lambda + c$.

### (b) Algebraic multiplicities are the same

Let $p(t) = \det(A - tI)$ be the characteristic polynomial of $A$.

Let $q(t)$ be the characteristic polynomial of the matrix $A + cI$ .

Then we have

$$\begin{aligned}q(t) &= \det\big((A + cI) - t\big) = \det\big(A - (t - c)\big) \\ &= p(t - c).\end{aligned}$$

Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of $A$ with algebraic multiplicities $n_1, \dots, n_k$, respectively.

Then we have

$$p(t) = \pm \prod_{i=1}^{k} (t - \lambda_i)^{n_i}.$$

It follows that we have

$$\begin{aligned} q(t) &= p(t-c) \\ &= \pm \prod_{i=1}^{k} (t-c-\lambda_i)^{n_i} \\ &= \pm \prod_{i=1}^{k} \left(t-(\lambda_i+c)\right)^{n_i}. \end{aligned}$$

From the last equation, we read that the eigenvalues of the matrix $A+cI$ are $\lambda_i+c$ with algebraic multiplicity $n_i$ for $i=1,\dots,k$.
Thus, geometric multiplicities of $\lambda$ and $\lambda-c$ are the same.

### (c) How about geometric multiplicities?

From part (a), we know that eigenvectors of $\lambda$ are eigenvectors of $\lambda-c$.
Reversing the argument, the eigenvectors of $\lambda+c$ are eigenvectors of $\lambda$.

Thus, the eigenvectors correspond one to one, and the eigenspace of $\lambda$ is the same as the eigenspace of $\lambda+c$.
Hence their geometric multiplicities are the same.

## Applications

As applications of this problem, consider the following problems.

*Problem*. Find all the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 9 & 9 & 9 \\ 9 & 3 & 9 & 9 \\ 9 & 9 & 3 & 9 \\ 9 & 9 & 9 & 3 \end{bmatrix}.$$

This is a problem of a Linear Algebra final exam at Harvard University.
For a solution of this problem, see the post "Eigenvalues and eigenvectors of matrix whose diagonal entries are 3 and 9 elsewhere".

*Problem*. Find the determinant of the following matrix

$$A = \begin{bmatrix} 6 & 2 & 2 & 2 & 2 \\ 2 & 6 & 2 & 2 & 2 \\ 2 & 2 & 6 & 2 & 2 \\ 2 & 2 & 2 & 6 & 2 \\ 2 & 2 & 2 & 2 & 6 \end{bmatrix}.$$

This is also a problem of a Linear Algebra final exam at Harvard University.

See the post "Determinant of matrix whose diagonal entries are 6 and 2 elsewhere" for a solution.

(28) Find the determinant of the following matrix

$$A = \begin{bmatrix} 6 & 2 & 2 & 2 & 2 \\ 2 & 6 & 2 & 2 & 2 \\ 2 & 2 & 6 & 2 & 2 \\ 2 & 2 & 2 & 6 & 2 \\ 2 & 2 & 2 & 2 & 6 \end{bmatrix}.$$

(*Harvard University*)

## Hint.

Computing the determinant directly by hand is tedious.

So use the fact that the determinant of a matrix $A$ is the product of all eigenvalues of $A$.

However, finding the eigenvalue of $A$ itself is as complicated as computing the determinant of $A$.

Instead, first determine the eigenvalues of $B = A - 4I$.

Then use the fact that if $\lambda$ is an eigenvalue of $B$, then $\lambda + 4$ is an eigenvalue of $A$.

For a proof of this fact, see the post "Eigenvalues and algebraic/geometric multiplicities of matrix $A + cI$ ".

## Solution.

Let $B = A - 4I$, where $I$ is the $5 \times 5$ identity matrix.

Then every entry of $B$ is $2$.

By elementary row operations, we can reduced the matrix $B$ into

$$B = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, the rank of $B$ is $1$, hence the nullity is $4$ by the rank-nullity theorem.

It follows that $0$ is an eigenvalue of $B$ and its geometric multiplicity is $4$.

Since all entries of $B$ are equal, we compute

$$B\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}=10\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}.$$

This yields that $10$ is an eigenvalue of $B$ and the vector $[1\,1\,1\,1\,1]^{\mathrm{T}}$ is an eigenvector corresponding to $10$.

Combining these observations, we see that the matrix $B$ has eigenvalues $0$ and $10$ with (algebraic) multiplicities $4$ and $0$, respectively.

Since $A=B+4I$, the eigenvalues of $A$ are $4$ and $14$ with algebraic multiplicities $4$ and $0$, respectively.
(See Problem Eigenvalues and algebraic/geometric multiplicities of matrix $A+cI$.)

The determinant of $A$ is the product of all the eigenvalues of $A$ (counting multiplicities). Thus we have

$$\det(A)=4^4\cdot 14=3584.$$

㉙ Find all eigenvalues of the matrix

$$A=\begin{bmatrix}0 & i & i & i\\ i & 0 & i & i\\ i & i & 0 & i\\ i & i & i & 0\end{bmatrix},$$

where $i=\sqrt{-1}$. For each eigenvalue of $A$, determine its algebraic multiplicity and geometric multiplicity.

### Proof.

We use an indirect method to find the eigenvalues of $A$ as follows.
Let $B=A+iI$. Then note that if $\lambda$ is an eigenvalue of $B$, then $\lambda-i$ is an eigenvalue of $A$.

Furthermore, the algebraic (respectively, geometric) multiplicity of $\lambda$ is the same as the algebraic (respectively, geometric) multiplicity of $\lambda-i$.

We reduced the matrix $B$ as follows:

$$B=\begin{bmatrix}i & i & i & i\\ i & i & i & i\\ i & i & i & i\\ i & i & i & i\end{bmatrix}\xrightarrow{\substack{R_2-R_1\\ R_3-R_1\\ R_4-R_1}}\begin{bmatrix}i & i & i & i\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\xrightarrow{\frac{1}{i}R_1}\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

It follows that the rank of $B$ is $1$, and thus the nullity of $B$ is $3$ by the rank-nullity theorem.

Since the nullity is greater than $0$, it yields that $\lambda = 0$ is an eigenvalue of $A$.

---

Recall that the eigenspace $E_0$ of $\lambda = 0$ is the null space of $B$.
Hence the geometric multiplicity of $\lambda = 0$ is the nullity of $B$, thus it is $3$.
It follows that the algebraic multiplicity of $\lambda = 0$ is either $3$ or $4$.

---

Recall that the sum of all eigenvalues is equal to the trace of $B$, which is $\operatorname{tr}(B) = 4i$.
Thus, we see that $\lambda = 0$ is not the only eigenvalue.

There can be only one more eigenvalue of $B$ since the sum of algebraic multiplicities of all eigenvalues of $B$ is $4$.
It follows that $4i$ is an eigenvalue of $B$ with algebraic multiplicity $1$, and hence the geometric multiplicity is also $1$.
It yields tha the algebraic multiplicity of $\lambda = 0$ must be $3$.

---

Another way to see that $4i$ is an eigenvalue of $B$ is to notice that the vector

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

is an eigenvector. In fact, we have

$$B \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 4i \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

(Since all the entries are the same, it might not so difficult to notice this.)

---

In summary so far, we have obtained that eigenvalues of $B$ are $0$ and $4i$ with the algebraic/geometric multiplicities are $3$ and $1$, respectively.

Since $A = B - iI$, it follows that the eigenvalues of $A$ are $0 - i = -i$ and $4i - i = 3i$ with the algebraic/geometric multiplicities are $3$ and $1$, respectively.

## Comment.

This is the second problem of Quiz 13 (Take Home Quiz) for Math 2568 (Introduction to Linear Algebra) at OSU in Spring 2017.
The first problem of Quiz 13 is “Quiz 13 (Part 1) Diagonalize a matrix.“.

30 Find all the eigenvalues and eigenvectors of the matrix

$$A=\begin{bmatrix} 10001 & 3 & 5 & 7 & 9 & 11 \\ 1 & 10003 & 5 & 7 & 9 & 11 \\ 1 & 3 & 10005 & 7 & 9 & 11 \\ 1 & 3 & 5 & 10007 & 9 & 11 \\ 1 & 3 & 5 & 7 & 10009 & 11 \\ 1 & 3 & 5 & 7 & 9 & 10011 \end{bmatrix}.$$

(*MIT*)

## Solution.

Let

$$B=A-10000I,$$

where $I$ is the $6\times 6$ identity matrix. That is, we have

$$B=\begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 11 \\ 1 & 3 & 5 & 7 & 9 & 11 \\ 1 & 3 & 5 & 7 & 9 & 11 \\ 1 & 3 & 5 & 7 & 9 & 11 \\ 1 & 3 & 5 & 7 & 9 & 11 \\ 1 & 3 & 5 & 7 & 9 & 11 \end{bmatrix}.$$

Since all the rows are the same, the matrix $B$ is singular and hence $\lambda=0$ is an eigenvalue of $B$. Let us determine the geometric multiplicity of $\lambda=0$ (namely, the dimension of the null space of $B$).

We apply elementary row operations to $B$ and obtain

$$B\xrightarrow{\text{elementary row operations}}\begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, if $B\mathrm{x}=0$, then we have

$$x_1=-3x_2-5x_3-7x_4-9x_5-11x_6.$$

It follows from this that basis vectors of the eigenspace $E_0=\mathcal{N}(B)$ are

$$\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} -5 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} -7 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} -9 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},\begin{bmatrix} -11 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

and hence the geometric multiplicity corresponding to $\lambda=0$ is $5$.

By inspection, we see that

$$B\mathrm{v} = 36\mathrm{v},$$

where

$$\mathrm{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Thus, it yields that $\lambda = 36$ is an eigenvalue of $B$ and $\mathrm{v}$ is a corresponding eigenvector.

---

Recall that the algebraic multiplicity of an eigenvalue is greater than or equal to the geometric multiplicity.

Also the sum of algebraic multiplicities of all eigenvalues of $B$ is equal to $6$ since $B$ is a $6 \times 6$ matrix.

It follows from this observation that we determine that the algebraic multiplicity of $\lambda = 0$ is $5$ and the algebraic and geometric multiplicities of $\lambda = 36$ are both $1$.
Hence the vector $\mathrm{v}$ form a basis of the eigenspace $E_{36}$.

---

Now that we have determined eigenvalues and eigenvectors of $B$, we can deduce those of $A$ as follows.

In general, if $A = B + cI$, then the eigenvalues of $A$ are $\lambda + c$, where $\lambda$ are eigenvalues of $B$.
The eigenvectors for $A$ corresponding to $\lambda + c$ are exactly the eigenvectors for $B$ corresponding $\lambda$.
(See the post "Eigenvalues and algebraic/geometric multiplicities of matrix $A + cI$" for a proof.)

---

In the current problem, we have $A = B + 10000I$, and thus $c = 10000$.
Therefore, the eigenvalues of $A$ are $10000, 10036$.
Eigenvectors corresponding to $10000$ are

$$x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -7 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -9 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -11 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

where $(x_2, x_3, x_4, x_5, x_6) \neq (0, 0, 0, 0, 0, 0)$.

The eigenvector corresponding to $10036$ is

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

where $a$ is any nonzero scalar.

31 Consider the matrix

$$A = \begin{bmatrix} 3/2 & 2 \\ -1 & -3/2 \end{bmatrix} \in M_{2\times 2}(\mathbb{R}).$$

**(a)** Find the eigenvalues and corresponding eigenvectors of $A$.

**(b)** Show that for $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbb{R}^2$, we can choose $n$ large enough so that the length $\|A^n\mathbf{v}\|$ is as small as we like.

(*University of California, Berkeley*)

## Proof.

### (a) Find the eigenvalues and corresponding eigenvectors of $A$ .

To find the eigenvalues of $A$ , we compute the characteristic polynomial $p(t)$ of $A$ .

We have

$$\begin{aligned} p(t) &= \det(A - tI) \\ &= \begin{vmatrix} 3/2 - t & 2 \\ -1 & -3/2 - t \end{vmatrix} \\ &= t^2 - 1/4. \end{aligned}$$

Since the eigenvalues are roots of the characteristic polynomials, the eigenvalues of $A$ are $\pm 1/2$.

Next we find the eigenvectors corresponding to eigenvalue $1/2$.

These are the solutions of $(A - \frac{1}{2}I)\mathrm{x} = 0$ .

We have

$$A - \frac{1}{2}I = \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix} \xrightarrow{R_2 + R_1} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Thus, the solution $\mathrm{x}$ satisfies $x_1 = -2x_2$, and the eigenvectors are

$$\mathrm{x} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

where $x_2$ is a nonzero scalar.

Similarly, we find the eigenvectors corresponding to the eigenvalue $-1/2$.

We solve $(A + \frac{1}{2}I)\mathrm{x} = 0$ .

We have

$$A + \frac{1}{2}I = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} \xrightarrow[\text{then } R_2 + R_1]{\frac{1}{2}R_1} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Thus, we have $x_1 = -x_2$, and the eigenvectors are

$$\mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

where $x_2$ is a nonzero scalar.

(b) We can choose $n$ large enough so that the length $\|A^n \mathbf{v}\|$ is as small as we like.

We express the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of eigenvectors $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ corresponding to eigenvalues $1/2$ and $-1/2$, respectively.

Let

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1 \begin{bmatrix} -2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

for some scalars $c_1, c_2$.

Solving this for $c_1, c_2$, we find that $c_1 = -1$ and $c_2 = 1$, and thus we have

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = - \begin{bmatrix} -2 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then for any positive integer $n$, we have

$$\begin{aligned} A^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= -A^n \begin{bmatrix} -2 \\ 1 \end{bmatrix} + A^n \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= -\left(\frac{1}{2}\right)^n \begin{bmatrix} -2 \\ 1 \end{bmatrix} + \left(-\frac{1}{2}\right)^n \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \left(\frac{1}{2}\right)^n \begin{bmatrix} 2-(-1)^n \\ -1+(-1)^n \end{bmatrix} \end{aligned}$$

Note that in the second equality we used the following fact: If $A\mathbf{x} = \lambda\mathbf{x}$, then $A^n\mathbf{x} = \lambda^n\mathbf{x}$.

Then the length is

$$\begin{aligned} \left\| A^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\| &= \left(\frac{1}{2}\right)^n \sqrt{(2-(-1)^n)^2 + (-1+(-1)^n)^2} \\ &\leq \left(\frac{1}{2}\right)^n \sqrt{3^2 + 2^2} \\ &= \sqrt{13} \left(\frac{1}{2}\right)^n \to 0 \text{ as } n \text{ tends to infinity.} \end{aligned}$$

Therefore, we can choose $n$ large enough so that the length $\|A^n \mathbf{v}\|$ is as small as we wish.

32 Let $F$ and $H$ be an $n \times n$ matrices satisfying the relation $HF - FH = -2F$.
**(a)** Find the trace of the matrix $F$.
**(b)** Let $\lambda$ be an eigenvalue of $H$ and let $\mathbf{v}$ be an eigenvector corresponding to $\lambda$. Show that there exists an positive integer $N$ such that $F^N\mathbf{v} = \mathbf{0}$.

Hint.

33 For (a), take the trace of the both sides of the given relation.

㉞ For (b), show that if $F^k\mathbf{v} \neq \mathbf{0}$ then there are infinitely many eigenvalues, hence a contradiction.

## Proof.

### (a) The trace of the matrix $F$

Using the given relation we compute the trace of $F$ as follows.
By taking the trace of both sides we have

$$\operatorname{tr}(-2F) = \operatorname{tr}(HF - FH).$$

The right hand side is $-2\operatorname{tr}(F)$ and the left hand side is

$$\begin{aligned}\operatorname{tr}(HF - FH) &= \operatorname{tr}(HF) - \operatorname{tr}(FH)\\ &= \operatorname{tr}(HF) - \operatorname{tr}(HF) = 0.\end{aligned}$$

Therefore we have $\operatorname{tr}(F) = 0$.

### (b) There exists an positive integer $N$ such that $F^N\mathbf{v} = \mathbf{0}$.

Since $\mathbf{v}$ is an eigenvector corresponding to the eigenvalue $\lambda$ of $H$, we have $H\mathbf{v} = \lambda\mathbf{v}$ or equivalently $(H - \lambda I)\mathbf{v} = \mathbf{0}$.

Now we compute

$$\begin{aligned}F(H - \lambda I) &= FH - \lambda F\\ &= (HF + 2F) - \lambda F = (H - (\lambda - 2)I)F.\end{aligned}$$

Therefore we have

$$F(H - \lambda I) = (H - (\lambda - 2)I)F.$$

Evaluating at $\mathbf{v}$, we obtain

$$\mathbf{0} = F(H - \lambda I)\mathbf{v} = (H - (\lambda - 2)I)F\mathbf{v}.$$

If $F\mathbf{v} \neq \mathbf{0}$, then this equality implies that $F\mathbf{v}$ is an eigenvector corresponding to the eigenvalue $\lambda - 2$ of $H$. In this case we further calculate

$$\begin{aligned}\mathbf{0} &= F(H - (\lambda - 2)I)F\mathbf{v}\\ &= (FH - (\lambda - 2)F)F = (HF + 2F - (\lambda - 2)F)F\mathbf{v}\\ &= (H - (\lambda - 4))F^2\mathbf{v}.\end{aligned}$$

If the vector $F^2\mathbf{v} \neq \mathbf{0}$, then this equality implies that $F^2\mathbf{v}$ is an eigenvector corresponding to the eigenvalue $\lambda - 4$ of $H$.
Repeating this procedure, we see that

$$\mathbf{0} = (H - (\lambda - 2k))F^k\mathbf{v}$$

for all $k$.

Therefore, if $F^k\mathbf{v}$ is nonzero vector for all $k$, then there are infinitely many eigenvalues $\lambda - 2k$ but this is impossible since $H$ is an $n \times n$ matrix and hence $H$ has at most $n$ eigenvalues. Therefore there exists $N$ such that $F^N\mathbf{v} = \mathbf{0}$.

35 Let $A$ and $B$ be $n \times n$ matrices and assume that they commute: $AB = BA$. Then prove that the matrices $A$ and $B$ share at least one common eigenvector.

**Proof.**

Let $\lambda$ be an eigenvalue of $A$ and let $\mathrm{x}$ be an eigenvector corresponding to $\lambda$. That is, we have $A\mathrm{x} = \lambda\mathrm{x}$ .
Then we claim that the vector $\mathrm{v} := B\mathrm{x}$ belongs to the eigenspace $E_\lambda$ of $\lambda$.
In fact, as $AB = BA$ we have

$$A\mathrm{v} = AB\mathrm{x} = BA\mathrm{x} = \lambda B\mathrm{x} = \lambda\mathrm{v}.$$

Hence $\mathrm{v} \in E_\lambda$ .

Now, let $\{\mathrm{x}_1, \ldots, \mathrm{x}_k\}$ be an eigenbasis of the eigenspace $E_\lambda$ .
Set $\mathrm{v}_i = B\mathrm{x}_i$ for $i = 1, \ldots, k$ .
The above claim yields that $\mathrm{v}_i \in E_\lambda$ , and hence we can write

$$\mathrm{v}_i = c_{1i}\mathrm{x}_1 + c_{2i}\mathrm{x}_2 + \cdots + c_{ki}\mathrm{x}_k \qquad (*)$$

for some scalars $c_{1i}, c_{2i}, \ldots, c_{ki}$ .

Extend the basis $\{\mathrm{x}_1, \ldots, \mathrm{x}_k\}$ of $E_\lambda$ to a basis

$$\{\mathrm{x}_1, \ldots, \mathrm{x}_k, \mathrm{x}_{k+1}, \ldots, \mathrm{x}_n\}$$

of $\mathbb{R}^n$ by adjoining vectors $\mathrm{x}_{k+1}, \ldots, \mathrm{x}_n$ .

Then we obtain using (*)

$$\begin{aligned} &B[\mathrm{x}_1, \ldots, \mathrm{x}_k, \mathrm{x}_{k+1}, \ldots, \mathrm{x}_n] \\ &= [B\mathrm{x}_1, \ldots, B\mathrm{x}_k, B\mathrm{x}_{k+1}, \ldots, B\mathrm{x}_n] \\ &= [\mathrm{v}_1, \ldots, \mathrm{v}_k, B\mathrm{x}_{k+1}, \ldots, B\mathrm{x}_n] \\ &= [\mathrm{x}_1, \ldots, \mathrm{x}_k, \mathrm{x}_{k+1}, \ldots, \mathrm{x}_n] \left[\begin{array}{c|c} C & D \\ \hline O & F \end{array}\right], \end{aligned} \qquad (**)$$

where $C = (c_{ij})$ is the $k \times k$ matrix whose entries are the coefficients $c_{ij}$ of the linear combination (*), $O$ is the $(n-k) \times k$ zero matrix, $D$ is a $k \times (n-k)$ matrix, and $F$ is an $(n-k) \times (n-k)$ matrix.

---

Let $P = [\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n]$.
As the column vectors of $P$ are linearly independent, $P$ is invertible.

From (**), we obtain

$$P^{-1}BP = \left[\begin{array}{c|c} C & D \\ \hline O & F \end{array}\right].$$

It follows that

$$\det(B - tI) = \det(P^{-1}BP - tI) = \left|\begin{array}{c|c} C - tI & D \\ \hline O & F - tI \end{array}\right| = \det(C - tI)\det(F - tI).$$

---

Let $\mu$ be an eigenvalue of the matrix $C$ and let $\mathbf{a}$ be an eigenvector corresponding to $\mu$.
Then as $\det(C - \mu I) = 0$, we see that $\det(B - \mu I) = 0$ and $\mu$ is an eigenvalue of $B$.

Write

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \neq 0$$

and define a new vector by

$$\mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k \in E_\lambda.$$

Then $\mathbf{y}$ is an eigenvector in $E_\lambda$ since it is a nonzero (as $\mathbf{a} \neq \mathbf{0}$) linear combination of the basis $E_\lambda$.

---

Multiplying $BP = P \left[\begin{array}{c|c} C & D \\ \hline O & F \end{array}\right]$ from (**) by the $n$-dimensional vector

$$\begin{bmatrix} \mathbf{a} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

on the right, we have

$$BP\begin{bmatrix}\mathrm{a}\\0\\\vdots\\0\end{bmatrix}=P\left[\begin{array}{c|c}C&D\\\hline O&F\end{array}\right]\begin{bmatrix}\mathrm{a}\\0\\\vdots\\0\end{bmatrix}.$$

---

The left hand side is equal to

$$B[a_1\mathrm{x}_1+\cdots+a_k\mathrm{x}_k]=B\mathrm{y}.$$

On the other hand, the right hand side is equal to

$$P\begin{bmatrix}C\mathrm{a}\\0\end{bmatrix}=P\begin{bmatrix}\mu\mathrm{a}\\0\end{bmatrix}=\mu P\begin{bmatrix}\mathrm{a}\\0\end{bmatrix}=\mu[a_1\mathrm{x}_1+\cdots+a_k\mathrm{x}_k]=\mu\mathrm{y}.$$

---

Therefore, we obtain

$$B\mathrm{y}=\mu\mathrm{y}.$$

This proves that the vector $\mathrm{y}$ is eigenvector of both $A$ and $B$.

Hence, this completes the proof.

(36) Let $a$ and $b$ be two distinct positive real numbers. Define matrices

$$A:=\begin{bmatrix}0&a\\a&0\end{bmatrix},\ B:=\begin{bmatrix}0&b\\b&0\end{bmatrix}.$$

Find all the pairs $(\lambda, X)$, where $\lambda$ is a real number and $X$ is a non-zero real matrix satisfying the relation

$$AX+XB=\lambda X.$$

(*The University of Tokyo*)

## Hint.

(37) Let $X=\begin{bmatrix}x_1&x_2\\x_3&x_4\end{bmatrix}$ and compute (*).

(38) Compare entries of matrices.

(39) Rewrite it as a matrix equation.

(40) The problem is now translated to a problem of eigenvalue/eigenvectors.

## Solution.

Let $X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$.

Then the relation (*) becomes

$$a \begin{bmatrix} x_3 & x_4 \\ x_1 & x_2 \end{bmatrix} + b \begin{bmatrix} x_2 & x_1 \\ x_4 & x_3 \end{bmatrix} = \lambda \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

Comparing the entries of matrices, we obtain four equations

$$\begin{aligned} ax_3 + bx_2 &= \lambda x_1 \\ ax_4 + bx_1 &= \lambda x_2 \\ ax_1 + bx_4 &= \lambda x_3 \\ ax_2 + bx_3 &= \lambda x_4. \end{aligned}$$

We rewrite these equations into the matrix equation

$$\begin{bmatrix} -\lambda & b & a & 0 \\ b & -\lambda & 0 & a \\ a & 0 & -\lambda & b \\ 0 & a & b & -\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus the problem is to find all the pairs $(\lambda, X)$ satisfying this matrix equation.

Let

$$C = \begin{bmatrix} 0 & b & a & 0 \\ b & 0 & 0 & a \\ a & 0 & 0 & b \\ 0 & a & b & 0 \end{bmatrix}.$$

Then the problem is equivalent to find all eigenvalues and eigenvectors of the matrix $C$.

So we first compute the characteristic polynomial of $C$ to find eigenvalues.

We have

$$\begin{aligned} \det(C - \lambda I) &= \begin{vmatrix} -\lambda & b & a & 0 \\ b & -\lambda & 0 & a \\ a & 0 & -\lambda & b \\ 0 & a & b & -\lambda \end{vmatrix} \\ &= \lambda^4 - 2(a^2 + b^2)\lambda^2 + (a^2 - b^2)^2. \end{aligned}$$

(Use cofactor expansion and simplify to obtain this.)

We solve

$$\lambda^4 - 2(a^2 + b^2)\lambda^2 + (a^2 - b^2)^2 = 0$$

for $\lambda$.

By the quadratic formula we have

$$\begin{aligned} \lambda^2 &= a^2 + b^2 \pm \sqrt{(a^2 + b^2)^2 - (a^2 - b^2)^2} \\ &= a^2 + b^2 \pm 2ab = (a \pm b)^2. \end{aligned}$$

Hence we obtained four eigenvalues $\lambda = \pm(a \pm b)$.
Note that since we have four distinct eigenvalues, each eigenspace is one dimensional.

Now, let us find eigenvectors. First consider the eigenvalue $\lambda = a + b$.
In this case,

$$C - \lambda I = \begin{bmatrix} -a-b & b & a & 0 \\ b & -a-b & 0 & a \\ a & 0 & -a-b & b \\ 0 & a & b & -a-b \end{bmatrix}.$$

The eigenvector is the solutions of $(C - \lambda I)\mathbf{x} = \mathbf{0}$. The typical method to find eigenvector is to reduce the matrix $C - \lambda I$ into row echelon form.
But as we noted earlier each eigenspace is one dimensional, so if we find one nonzero vector solution, then it is a basis of the eigenspace.

For this specific matrix, we see that $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ satisfies $(C - \lambda I)\mathbf{x} = \mathbf{0}$. Therefore all the eigenvectors corresponding to eigenvalue $\lambda = a + b$ is

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} t$$

for any nonzero scalar $t$.

Similarly, we find that

$$\begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} t, \quad \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} t, \quad \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} t$$

for any nonzero $t$ are all the eigenvector corresponding to eigenvalue $\lambda = -a-b, a-b, -a+b$, respectively.

Therefore the solution to the problem is that the pair $(\lambda, X)$ is one of the following paris.

$$\left(a+b, t\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right), \left(-a-b, t\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\right),$$
$$\left(a-b, t\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\right), \left(-a+b, t\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\right)$$

for nonzero $t$ (because $X$ must be a nonzero matrix.)

# Diagonalization of Matrices

The general procedure of the diagonalization is explained in the post "How to Diagonalize a Matrix. Step by Step Explanation".

**Definition**

Let $A, B$ be $n \times n$ matrices.

1. $A$ and $B$ are **similar** if there exists a nonsingular matrix $P$ such that $P^{-1}AP = B$.
2. $A$ is **diagonalizable** if there exist a diagonal matrix $D$ and nonsingular matrix $P$ such that $P^{-1}AP = D$. (Namely, if $A$ is diagonalizable if it is similar to a diagonal matrix.)
3. $A$ is said to be **defective** if there is an eigenvalue $\lambda$ of $A$ such that the geometric multiplicity of $\lambda$ is less than the algebraic multiplicity of $\lambda$.

**Summary**

Let $A, B$ be $n \times n$ matrices.

1. If $A$ and $B$ are similar, then the characteristic polynomials of $A$ and $B$ are the same. Hence the eigenvalues of $A, B$ and their algebraic multiplicities are the same.
2. $A$ is diagonalizable if and only if $A$ is not defective.
3. $A$ is diagonalizable if and only if $\mathbb{R}^n$ has an eigenbasis of $A$ (a basis consisting of eigenvectors).
4. $A$ is diagonalizable if and only if there are $n$ linearly independent eigenvectors of $A$.
5. If $A$ has $n$ distinct eigenvalues, then $A$ is diagonalizable.
6. If $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent eigenvectors of $A$ corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$ (not necessarily distinct), then $S^{-1}AS = D$, where $S = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

## Problems

1. Let $A, B$, and $C$ be $n \times n$ matrices and $I$ be the $n \times n$ identity matrix. Prove the following statements.
   **(a)** If $A$ is similar to $B$, then $B$ is similar to $A$.
   **(b)** $A$ is similar to itself.
   **(c)** If $A$ is similar to $B$ and $B$ is similar to $C$, then $A$ is similar to $C$.
   **(d)** If $A$ is similar to the identity matrix $I$, then $A = I$.
   **(e)** If $A$ or $B$ is nonsingular, then $AB$ is similar to $BA$.
   **(f)** If $A$ is similar to $B$, then $A^k$ is similar to $B^k$ for any positive integer $k$.

### Definition.

We say that a matrix $A$ is **similar** to a matrix $B$ is there exists a nonsingular (invertible) matrix $P$ such that

$$B = P^{-1}AP.$$

Proof.

(a) If $A$ is similar to $B$, then $B$ is similar to $A$.

If $A$ is similar to $B$, then there exists a nonsingular matrix $P$ such that $B = P^{-1}AP$.
Let $Q = P^{-1}$. Since $P$ is nonsingular, so is $Q$.
Then we have

$$\begin{aligned} Q^{-1}BQ &= (P^{-1})^{-1}BP^{-1} = PBP^{-1} \\ &= P(P^{-1}AP)P^{-1} = IAI = A. \end{aligned}$$

Hence $B$ is similar to $A$.

(b) $A$ is similar to iteself.

Since the identity matrix $I$ is nonsingular and we have

$$A = I^{-1}AI,$$

the matrix $A$ is similar to $A$ itself.

(c) If $A$ is similar to $B$ and $B$ is similar to $C$, then $A$ is similar to $C$.

If $A$ is similar to $B$, we have

$$B = P^{-1}AP,$$

for some nonsingular matrix $P$.
Also, if $B$ is similar to $C$, we have

$$C = Q^{-1}BQ,$$

for some nonsingular matrix $Q$.
Then we have

$$\begin{aligned} C &= Q^{-1}BQ \\ &= Q^{-1}(P^{-1}AP)Q \\ &= (PQ)^{-1}A(PQ). \end{aligned}$$

Let $R = PQ$. Since both $P$ and $Q$ are nonsingular, $R = PQ$ is also nonsingular.
The above computation yields that we have

$$C = R^{-1}AR,$$

hence $A$ is similar to $C$.

(d) If $A$ is similar to the identity matrix $I$, then $A = I$.

Since $A$ is similar to $I$, there exists a nonsingular matrix $P$ such that

$$A = P^{-1}IP.$$

Since $P^{-1}IP = I$, we have $A = I$.

(e) If $A$ or $B$ is nonsingular, then $AB$ is similar to $BA$.

Suppose first that $A$ is nonsingular. Then $A$ is invertible, hence the inverse matrix $A^{-1}$ exists.
Then we have

$$A^{-1}(AB)A = A^{-1}ABA = IBA = BA,$$

hence $AB$ and $BA$ are similar.

Analogously, if $B$ is nonsingular, then the inverse matrix $B^{-1}$ exists.
We have

$$B^{-1}(BA)B = B^{-1}BAB = IAB = AB,$$

hence $AB$ and $BA$ are similar.

(f) If $A$ is similar to $B$, then $A^k$ is similar to $B^k$.

If $A$ is similar to $B$, then we have

$$B = P^{-1}AP$$

for some nonsingular matrix $P$.
Then we have for a positive integer $k$

$$\begin{aligned} B^k &= (P^{-1}AP)^k \\ &= \underbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)}_{k \text{ times}} \\ &= P^{-1}A^kP \end{aligned}$$

since we can cancel $P$ and $P^{-1}$ in between.
Hence $A^k$ and $B^k$ are similar.

## Comment.

Part (a), (b), (c) show that similarity is an equivalence relation.

**2** Prove that if $A$ and $B$ are similar matrices, then their determinants are the same.

## Proof.

Suppose that $A$ and $B$ are similar. Then there exists a nonsingular matrix $S$ such that

$$S^{-1}AS = B$$

by definition.

Then we have

$$\begin{aligned}&\det(B)\\&=\det(S^{-1}AS)\\&=\det(S)^{-1}\det(A)\det(S)\\&\text{(by multiplicative properties of determinants)}\\&=\det(A)\\&\text{(since determinants are just numbers, hence commutative).}\end{aligned}$$

More generally, we can prove that if $A$ and $B$ are similar, then their characteristic polynomials are the same.

From this, we also can deduce that the determinants of $A$ and $B$ are the same as well as their traces are the same.

③ **(a)** Is the matrix $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ similar to the matrix $B = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$?

**(b)** Is the matrix $A = \begin{bmatrix} 0 & 1 \\ 5 & 3 \end{bmatrix}$ similar to the matrix $B = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$?

**(c)** Is the matrix $A = \begin{bmatrix} -1 & 6 \\ -2 & 6 \end{bmatrix}$ similar to the matrix $B = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$?

**(d)** Is the matrix $A = \begin{bmatrix} -1 & 6 \\ -2 & 6 \end{bmatrix}$ similar to the matrix $B = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$?

④ Determine whether the matrix $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ is diagonalizable. If so, find a nonsingular matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.

(*The Ohio State University*)

## Solution.

To determine whether the matrix $A$ is diagonalizable, we first find eigenvalues of $A$.

To do so, we compute the characteristic polynomial $p(t)$ of $A$:

$$p(t) = \begin{vmatrix} 1-t & 4 \\ 2 & 3-t \end{vmatrix} = (1-t)(3-t) - 8$$

$$= t^2 - 4t - 5 = (t+1)(t-5).$$

The roots of the characteristic polynomial $p(t)$ are eigenvalues of $A$.

Hence the eigenvalues of $A$ are $-1$ and $5$.

Since the $2 \times 2$ matrix $A$ has two distinct eigenvalues, it is diagonalizable.

To find the invertible matrix $S$, we need eigenvectors.

Let us find the eigenvectors corresponding to the eigenvalue $-1$.
By elementary row operations, we have

$$\begin{aligned} A-(-1)I = A+I &= \begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix} \\ &\xrightarrow{R_2-R_1} \begin{bmatrix} 2 & 4 \\ 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

It follows that the eigenvectors corresponding to $-1$ are of the form

$$a\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

for any nonzero scalar $a$.

---

Similarly, we have

$$A-5I = \begin{bmatrix} -4 & 4 \\ 2 & -2 \end{bmatrix} \xrightarrow{\frac{-1}{4}R_1} \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \xrightarrow{R_2-2R_1} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

Hence the eigenvectors corresponding to $5$ are of the form

$$b\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

for any nonzero scalar $b$.

Thus, $\mathbf{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are basis vectors of eigenspace $E_{-1}, E_5$, respectively.

---

Define

$$S := [\mathbf{u} \quad \mathbf{v}] = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then by the general procedure of the diagonalization, we have

$$S^{-1}AS = D,$$

where

$$D := \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}.$$

(5) Diagonalize the $2\times 2$ matrix $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ by finding a nonsingular matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.

## Solution.

The characteristic polynomial $p(t)$ of the matrix $A$ is

$$\begin{aligned} p(t) = \det(A - tI) &= \begin{vmatrix} 2-t & -1 \\ -1 & 2-1 \end{vmatrix} \\ &= (2-t)^2 - 1 = t^2 - 4t + 3 \\ &= (t-1)(t-3). \end{aligned}$$

It follows that the eigenvalues of $A$ are $\lambda = 1, 3$ with algebraic multiplicities are both $1$.
Hence, the geometric multiplicities are $1$ and thus any nonzero vector in eahc eigenspace forms a eigenbasis.

---

Now let us find a eigenbasis for each eigenspace $E_\lambda = \mathcal{N}(A - \lambda I)$.
For the eigenvalue $1$, we have

$$A - I = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \xrightarrow{R_2+R_1} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

This yields that the eigenvectors corresponding to the eigenvalue $1$ are $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with $x_2 \neq 0$. Hence

$$\mathrm{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in E_1$$

is an eigenbasis for $E_1$.

---

Similarly, as we have

$$A - 3I = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \xrightarrow{-R_1} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \xrightarrow{R_2+R_1} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},$$

we see that

$$\mathrm{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \in E_3$$

is an eigenbasis for $E_3$.

---

Let

$$S = [\mathrm{v}_1 \quad \mathrm{v}_2] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

Then the diagonalization procedure yields that $S$ is nonsingular and $S^{-1}AS = D$.

⑥ Diagonalize the matrix

$$A = \begin{bmatrix} 4 & -3 & -3 \\ 3 & -2 & -3 \\ -1 & 1 & 2 \end{bmatrix}$$

by finding a nonsingular matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.

## Diagonalization Procedure

Let $A$ be the $n \times n$ matrix that you want to diagonalize (if possible).

⑦ Find the characteristic polynomial $p(t)$ of $A$.

⑧ Find eigenvalues $\lambda$ of the matrix $A$ and their algebraic multiplicities from the characteristic polynomial $p(t)$.

⑨ For each eigenvalue $\lambda$ of $A$, find a basis of the eigenspace $E_\lambda$.
If there is an eigenvalue $\lambda$ such that the geometric multiplicity of $\lambda$, $\dim(E_\lambda)$, is less than the algebraic multiplicity of $\lambda$, then the matrix $A$ is not diagonalizable. If not, $A$ is diagonalizable, and proceed to the next step.

⑩ If we combine all basis vectors for all eigenspaces, we obtained $n$ linearly independent eigenvectors $\mathrm{v}_1, \mathrm{v}_2, \dots, \mathrm{v}_n$.

⑪ Define the nonsingular matrix

$$S = [\mathrm{v}_1 \mathrm{v}_2 \dots \mathrm{v}_n].$$

⑫ Define the diagonal matrix $D$, whose $(i, i)$-entry is the eigenvalue $\lambda$ such that the $i$-th column vector $\mathrm{v}_i$ is in the eigenspace $E_\lambda$.

⑬ Then the matrix $A$ is diagonalized as

$$S^{-1}AS = D.$$

## Example of a matrix diagonalization

Now let us examine these steps with an example.

Let us consider the following $3 \times 3$ matrix.

$$A = \begin{bmatrix} 4 & -3 & -3 \\ 3 & -2 & -3 \\ -1 & 1 & 2 \end{bmatrix}.$$

We want to diagonalize the matrix if possible.

### Step 1: Find the characteristic polynomial

The characteristic polynomial $p(t)$ of $A$ is

$$p(t) = \det(A - tI) = \begin{vmatrix} 4-t & -3 & -3 \\ 3 & -2-t & -3 \\ -1 & 1 & 2-t \end{vmatrix}.$$

Using the cofactor expansion, we get

$$p(t) = -(t-1)^2(t-2).$$

### Step 2: Find the eigenvalues

From the characteristic polynomial obtained in Step 1, we see that eigenvalues are

$$\lambda = 1 \text{ with algebraic multiplicity } 2$$

and

$$\lambda = 2 \text{ with algebraic multiplicity } 1.$$

### Step 3: Find the eigenspaces

Let us first find the eigenspace $E_1$ corresponding to the eigenvalue $\lambda = 1$.
By definition, $E_1$ is the null space of the matrix

$$A - I = \begin{bmatrix} 3 & -3 & -3 \\ 3 & -3 & -3 \\ -1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

by elementary row operations.
Hence if $(A - I)\mathbf{x} = 0$ for $\mathbf{x} \in \mathbb{R}^3$, we have

$$x_1 = x_2 + x_3.$$

Therefore, we have

$$E_1 = \mathcal{N}(A - I) = \left\{ \quad \mathbf{x} \in \mathbb{R}^3 \quad \middle| \quad \mathbf{x} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \right\}.$$

From this, we see that the set

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis for the eigenspace $E_1$.

Thus, the dimension of $E_1$, which is the geometric multiplicity of $\lambda = 1$, is $2$.

Similarly, we find a basis of the eigenspace $E_2 = \mathcal{N}(A - 2I)$ for the eigenvalue $\lambda = 2$.

We have

$$A - 2I = \begin{bmatrix} 2 & -3 & -3 \\ 3 & -4 & -3 \\ -1 & 1 & 0 \end{bmatrix} \rightarrow \cdots \rightarrow \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

by elementary row operations.

Then if $(A - 2I)\mathbf{x} = 0$ for $\mathbf{x} \in \mathbb{R}^3$, then we have

$$x_1 = -3x_3 \text{ and } x_2 = -3x_3.$$

Therefore we obtain

$$E_2 = \mathcal{N}(A - 2I) = \left\{ \mathbf{x} \in \mathbb{R}^3 \;\middle|\; \mathbf{x} = x_3 \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix} \right\}.$$

From this we see that the set

$$\left\{ \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix} \right\}$$

is a basis for the eigenspace $E_2$ and the geometric multiplicity is $1$.

Since for both eigenvalues, the geometric multiplicity is equal to the algebraic multiplicity, the matrix $A$ is not defective, and hence diagonalizable.

### Step 4: Determine linearly independent eigenvectors

From Step 3, the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix}$$

are linearly independent eigenvectors.

### Step 5: Define the invertible matrix $S$

Define the matrix $S = [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3]$. Thus we have

$$S = \begin{bmatrix} 1 & 1 & -3 \\ 1 & 0 & -3 \\ 0 & 1 & 1 \end{bmatrix}$$

and the matrix $S$ is nonsingular (since the column vectors are linearly independent).

### Step 6: Define the diagonal matrix $D$

Define the diagonal matrix

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Note that $(1,1)$-entry of $D$ is $1$ because the first column vector $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ of $S$ is in the eigenspace $E_1$, that is, $\mathbf{v}_1$ is an eigenvector corresponding to eigenvalue $\lambda = 1$.

Similarly, the $(2,2)$-entry of $D$ is $1$ because the second column $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ of $S$ is in $E_1$.

The $(3,3)$-entry of $D$ is $2$ because the third column vector $\mathbf{v}_3 = \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix}$ of $S$ is in $E_2$.

(The order you arrange the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ to form $S$ does not matter but once you made $S$, then the order of the diagonal entries is determined by $S$, that is, the order of eigenvectors in $S$.)

### Step 7: Finish the diagonalization

Finally, we can diagonalize the matrix $A$ as

$$S^{-1}AS = D,$$

where

$$S = \begin{bmatrix} 1 & 1 & -3 \\ 1 & 0 & -3 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

(Here you don't have to find the inverse matrix $S^{-1}$ unless you are asked to do so.)

## Diagonalization Problems and Examples

Check out the following problems about the diagonalization of a matrix to see if you understand the procedure.

**Problem**. Diagonalize

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

and compute $A^{100}$.

For a solution of this problem and related questions, see the post "Diagonalize a 2 by 2 Matrix $A$ and Calculate the Power $A^{100}$".

**Problem**. Determine whether the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

is diagonalizable. If it is diagonalizable, then find the invertible matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.

For a solution, check out the post "Diagonalize the 3 by 3 Matrix if it is Diagonalizable".

**Problem**. Let

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Determine whether the matrix $A$ is diagonalizable. If it is diagonalizable, then diagonalize $A$.

For a solution, see the post "Quiz 13 (Part 1) Diagonalize a matrix.".

**Problem**. Diagonalize the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

In the solution given in the post "Diagonalize the 3 by 3 Matrix Whose Entries are All One", we use an indirect method to find eigenvalues and eigenvectors.

The next problem is a diagonalization problem of a matrix with variables.

**Problem**.
Diagonalize the complex matrix

$$A = \begin{bmatrix} a & b-a \\ 0 & b \end{bmatrix}.$$

Using the result of the diagonalization, compute $A^k$ for each $k \in \mathbb{N}$.

The solution is given in the post↴
Diagonalize the Upper Triangular Matrix and Find the Power of the Matrix

## A Hermitian Matrix can be diagonalized by a unitary matrix

**Theorem**. If $A$ is a Hermitian matrix, then $A$ can be diagonalized by a unitary matrix $U$.

This means that there exists a unitary matrix $U$ such that $U^{-1}AU$ is a diagonal matrix.

**Problem**.
Diagonalize the Hermitian matrix

$$A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$

by a unitary matrix.

The solution is given in the post ↴
Diagonalize the $2 \times 2$ Hermitian Matrix by a Unitary Matrix

## More diagonalization problems

More Problems related to the diagonalization of a matrix are gathered in the following page:

Diagonalization of Matrices

⑭ Suppose that $A$ and $P$ are $3 \times 3$ matrices and $P$ is invertible matrix.
If

$$P^{-1}AP = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix},$$

then find all the eigenvalues of the matrix $A^2$.

## Proof (short version).

Let $B = P^{-1}AP$ . Since $B$ is an upper triangular matrix, its eigenvalues are diagonal entries $1, 4, 6$.

Since $A$ and $B = P^{-1}AP$ have the same eigenvalues, the eigenvalues of $A$ are $1, 4, 6$. Note that these are all the eigenvalues of $A$ since $A$ is a $3 \times 3$ matrix.

It follows that all the eigenvalues of $A^2$ are $1, 4^2, 6^2$, that is, $1, 16, 36$.

## Proof (long version.)

Let us put $B := P^{-1}AP$ .
The eigenvalues of $B$ are $1, 4, 6$ since $B$ is an upper triangular matrix and eigenvalues of an upper triangular matrix are diagonal entries.
We claim that the eigenvalues of $A$ and $B$ are the same.

To prove this claim, we show that their characteristic polynomials are equal.
Let $p_A(t), p_B(t)$ be the characteristic polynomials of $A, B$, respectively.
Then we have

$$\begin{aligned} p_B(t) &= \det(B - tI) = \det(P^{-1}AP - tI) \\ &= \det(P^{-1}(A - tI)P) \\ &= \det(P^{-1}) \det(A - tI) \det(P) \\ &= \det(P)^{-1} p_A(t) \det(P) \\ &= p_A(t). \end{aligned}$$

---

Therefore, the characteristic polynomials of $A$ and $B$ are the same, and hence the matrices $A$ and $B$ have the same eigenvalues.
Thus $1, 4, 6$ are eigenvalues of the matrix $A$ and these are all the eigenvalues of $A$ since a $3 \times 3$ matrix has at most three eigenvalues.

---

Now we claim that in general if $\lambda$ is an eigenvalue of $A$, then $\lambda^2$ is an eigenvalue of $A^2$. We prove this statement below. Assuming this claim for the moment, we finish the problem.
By the claim, the matrix $A^2$ has eigenvalues $1^2, 4^2, 6^2$. Thus all the eigenvalues of $A^2$ are

$$1, 16, 36.$$

---

To complete the solution, let us prove the last claim.
Since $\lambda$ is an eigenvalue of $A$, we have an eigenvector $\mathbf{x}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Multiplying by $A$ we obtain

$$\begin{aligned} A^2 T\mathbf{x} &= \lambda(A\mathbf{x}) \\ &= \lambda(\lambda\mathbf{x}) \\ &= \lambda^2\mathbf{x} \end{aligned}$$

and thus $\lambda^2$ is an eigenvalue of $A^2$.

15 Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Compute $A^n$ for any $n \in \mathbb{N}$.

## Plan.

We diagonalize the matrix $A$ and use this Problem.

## Steps.

⑯ Find eigenvalues and eigenvectors of the matrix $A$.

⑰ Diagonalize the matrix $A$.

⑱ Use the result of this Problem.

### Proof.

We first diagonalize the matrix $A$. We solve

$$\begin{aligned}\det(A-\lambda I) &= \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)^2-4=\lambda^2-2\lambda-3=(\lambda+1)(\lambda-3)=0\end{aligned}$$

and obtain the eigenvalues $\lambda=-1,3$.

To find an eigenvector $\mathbf{x}$ corresponding to $\lambda=-1$, we solve $(A+I)\mathbf{x}=0$ or

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We obtain an eigenvector $\mathbf{x}=\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ corresponding to $\lambda=-1$.

Similarly, solving $(A-3I)\mathbf{y}=0$, we obtain an eigenvector $\mathbf{y}=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ corresponding to $\lambda=3$.

Thus the invertible matrix $S=[\mathbf{x}\,\mathbf{y}]=\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ diagonalizes the matrix $A$, that is,

$$S^{-1}AS=\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \text{ or equivalently } A=S\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}S^{-1}.$$

Then for each $n\in\mathbb{N}$, we have

$$\begin{aligned}A^n &= \left(S\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}S^{-1}\right)^n = S\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}^n S^{-1} = S\begin{bmatrix} (-1)^n & 0 \\ 0 & 3^n \end{bmatrix}S^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} (-1)^n & 0 \\ 0 & 3^n \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} (-1)^n+3^n & (-1)^{n+1}+3^n \\ (-1)^{n+1}+3^n & (-1)^n+3^n \end{bmatrix}.\end{aligned}$$

(See this problem for the details of these computations.)

Therefore we obtained the formula

$$A^n=\frac{1}{2}\begin{bmatrix} (-1)^n+3^n & (-1)^{n+1}+3^n \\ (-1)^{n+1}+3^n & (-1)^n+3^n \end{bmatrix}.$$

### Comment.

Another typical method to compute a power of a square matrix is mathematical induction. To use it, we need to first compute several small powers like $A^2$ and $A^3$ and guess the formula for $A^n$.

If you can guess the formula, then the mathematical induction part is not difficult. But for this specific problem, the formula is a bit complicated to guess as you can see from the solution above. Thus we used diagonalization trick.

⑲ Let $A, B$ be matrices. Show that if $A$ is diagonalizable and if $B$ is similar to $A$, then $B$ is diagonalizable.

## Definitions/Hint.

Recall the relevant definitions.

⑳ Two matrices $A$ and $B$ are **similar** if there exists a nonsingular (invertible) matrix $S$ such that

$$S^{-1}BS = A.$$

㉑ A matrix $A$ is **diagonalizable** if $A$ is similar to a diagonal matrix. Namely, $A$ is diagonalizable if there exist a nonsingular matrix $S$ and a diagonal matrix $D$ such that

$$S^{-1}AS = D.$$

Some useful facts are

㉒ If $S$ and $T$ are invertible matrices, then we have

$$(TS)^{-1} = S^{-1}T^{-1}.$$

(Note the order of the product.)

㉓ A matrix is nonsingular if and only if its determinant is nonzero.

## Proof.

Since the matrix $A$ is diagonalizable, there exist a nonsingular matrix $S$ and a diagonal matrix $D$ such that

$$S^{-1}AS = D. \tag{*}$$

Also, since $B$ is similar to $A$, there exist a nonsingular matrix $T$ such that

$$T^{-1}BT = A. \tag{**}$$

Inserting the expression of $A$ from (**) into the equality (*), we obtain

$$\begin{aligned} D &= S^{-1}(T^{-1}BT)S \\ &= (S^{-1}T^{-1})B(TS) \\ &= (TS)^{-1}B(TS). \end{aligned} \tag{***}$$

---

Now let us put $U := TS$. Then the matrix $U$ is nonsingular.

(This is because we have

$$\det(U) = \det(TS) = \det(T)\det(S) \neq 0$$

since $T$ and $S$ are nonsingular matrices, hence their determinants are not zero.)

Therefore from (***) we have

$$D = U^{-1}BU,$$

where $D$ is a diagonal matrix and $U$ is a nonsingular matrix.

Thus $B$ is a diagonalizable matrix.

24 **(a)** Is every diagonalizable matrix invertible?
**(b)** Is every invertible matrix diagonalizable?

## Solution.

The answer is No.

### Counterexample

We give a counterexample. Consider the $2 \times 2$ zero matrix.

The zero matrix is a diagonal matrix, and thus it is diagonalizable.

However, the zero matrix is not invertible as its determinant is zero.

### More Theoretical Explanation

Let us give a more theoretical explanation.

If an $n \times n$ matrix $A$ is diagonalizable, then there exists an invertible matrix $P$ such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

where $\lambda_1, \dots, \lambda_n$ are eigenvalues of $A$.

Then we consider the determinants of the matrices of both sides.

The determinant of the left hand side is

$$\det(P^{-1}AP) = \det(P)^{-1}\det(A)\det(P) = \det(A).$$

On the other hand, the determinant of the right hand side is the product

$$\lambda_1\lambda_2\cdots\lambda_n$$

since the right matrix is diagonal.
Hence we obtain

$$\det(A)=\lambda_1\lambda_2\cdots\lambda_n.$$

(Note that it is always true that the determinant of a matrix is the product of its eigenvalues regardless diagonalizability.
See the post "Determinant/trace and eigenvalues of a matrix".)

Hence if one of the eigenvalues of $A$ is zero, then the determinant of $A$ is zero, and hence $A$ is not invertible.

The true statement is:

> a diagonal matrix is invertible if and only if its eigenvalues are nonzero.

## Is Every Invertible Matrix Diagonalizable?

Note that it is not true that every invertible matrix is diagonalizable.

For example, consider the matrix

$$A=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The determinant of $A$ is $1$, hence $A$ is invertible.
The characteristic polynomial of $A$ is

$$p(t)=\det(A-tI)=\begin{vmatrix} 1-t & 1 \\ 0 & 1-t \end{vmatrix}=(1-t)^2.$$

Thus, the eigenvalue of $A$ is $1$ with algebraic multiplicity $2$.
We have

$$A-I=\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

and thus eigenvectors corresponding to the eigenvalue $1$ are

$$a\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

for any nonzero scalar $a$.
Thus, the geometric multiplicity of the eigenvalue $1$ is $1$.
Since the geometric multiplicity is strictly less than the algebraic multiplicity, the matrix $A$ is defective and not diagonalizable.

## Is There a Matrix that is Not Diagonalizable and Not Invertible?

Finally, note that there is a matrix which is not diagonalizable and not invertible.

For example, the matrix $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is such a matrix.

## Summary

There are all possibilities.

25 Diagonalizable, but not invertible.
Example:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

26 Invertible, but not diagonalizable.
Example:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

27 Not diagonalizable and Not invertible.
Example:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

28 Diagonalizable and invertible
Example:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

29

Determine whether the matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is diagonalizable. If it is diagonalizable, then find the invertible matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.

## How to diagonalize matrices.

For a general procedure of the diagonalization of a matrix, please read the post “How to Diagonalize a Matrix. Step by Step Explanation“.

### Solution.

We first determine the eigenvalues of the matrix $A$.

To do so, we compute the characteristic polynomial $p(t)$ of $A$.

We have

$$\begin{aligned} p(t) &= \det(A - tI) \\ &= \begin{vmatrix} -t & 1 & 0 \\ -1 & -t & 0 \\ 0 & 0 & 2-t \end{vmatrix} \\ &= (-1)^{3+3}(2-t)\begin{vmatrix} -t & 1 \\ -1 & -t \end{vmatrix} && \text{by the third row cofactor expansion} \\ &= (2-t)(t^2+1). \end{aligned}$$

Thus the eigenvalues of $A$ are $2, \pm i$.

Since the $3 \times 3$ matrix $A$ has three distinct eigenvalues, it is diagonalizable.

To diagonalize $A$, we now find eigenvectors.

For the eigenvalue $2$, we compute

$$\begin{aligned} A - 2I = \begin{bmatrix} -2 & 1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} &\xrightarrow{-R_2} \begin{bmatrix} -2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} &\xrightarrow{R_2 + 2R_1} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \xrightarrow{\frac{1}{5}R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} &\xrightarrow{R_1 - 2R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Thus, the solutions $\mathbf{x}$ of $(A - 2I) = 0$ satisfy $x = y = 0$.

Hence the eigenspace is

$$E_2 = \mathcal{N}(A - 2I) = \operatorname{Span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

For the eigenvalue $i$, we compute

$$\begin{aligned} A - iI = \begin{bmatrix} -i & 1 & 0 \\ -1 & -i & 0 \\ 0 & 0 & 2-i \end{bmatrix} &\xrightarrow{iR_1} \begin{bmatrix} 1 & i & 0 \\ -1 & -i & 0 \\ 0 & 0 & 2-i \end{bmatrix} \\ \xrightarrow[\frac{1}{2-i}R_3]{R_2 + R_1} \begin{bmatrix} 1 & i & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} &\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & i & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

So the solutions $\mathbf{x}$ of $(A-iI)\mathbf{x}=0$ satisfy

$$x=-iy \text{ and } z=0.$$

Thus, the eigenspace is

$$E_i=\mathcal{N}(A-iI)=\operatorname{Span}\left\{\begin{bmatrix}1\\i\\0\end{bmatrix}\right\}.$$

---

Since $i$ and $-i$ are complex conjugate, their eigenspaces are also complex conjugate.

Hence the eigenspace for $-i$ is

$$E_{-i}=\operatorname{Span}\left\{\begin{bmatrix}1\\-i\\0\end{bmatrix}\right\}.$$

---

From these computations, we have obtained eigenvalues $2, i, -i$ and eigenvector corresponding to these are

$$\mathbf{v}_2=\begin{bmatrix}0\\0\\1\end{bmatrix}, \mathbf{v}_i=\begin{bmatrix}1\\i\\0\end{bmatrix}, \mathbf{v}_{-i}=\begin{bmatrix}1\\-i\\0\end{bmatrix}.$$

Let

$$S=[\mathbf{v}_2 \quad \mathbf{v}_i \quad \mathbf{v}_{-i}]=\begin{bmatrix}0&1&1\\0&i&-i\\1&0&0\end{bmatrix}$$

and

$$D=\begin{bmatrix}2&0&0\\0&i&0\\0&0&-i\end{bmatrix}.$$

Then $S$ is invertible and we have $S^{-1}AS=D$ by the diagonalization process.

30

For which values of constants $a, b$ and $c$ is the matrix $A=\begin{bmatrix}7&a&b\\0&2&c\\0&0&3\end{bmatrix}$ diagonalizable?

(*The Ohio State University*)

## Solution.

Note that the matrix $A$ is an upper triangular matrix.

Hence the eigenvalues of $A$ are diagonal entries $7, 2, 3$.

So the $3 \times 3$ matrix $A$ has three distinct eigenvalues.

This implies that $A$ is diagonalizable.

Hence, regardless of the values of $a, b, c$, the matrix $A$ is always diagonalizable.

Thus, $a, b, c$ can take arbitrary values.

(31) Let

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

For this problem, you may use the fact that both matrices have the same characteristic polynomial:

$$p_A(\lambda) = p_B(\lambda) = -(\lambda - 1)(\lambda + 2)^2.$$

**(a)** Find all eigenvectors of $A$.
**(b)** Find all eigenvectors of $B$.
**(c)** Which matrix $A$ or $B$ is diagonalizable?
**(d)** Diagonalize the matrix stated in (c), i.e., find an invertible matrix $P$ and a diagonal matrix $D$ such that $A = PDP^{-1}$ or $B = PDP^{-1}$.
(*Stanford University*)

## Hint.

See the post how to diagonalize a matrix for a review of the diagonalization process.

## Solution.

### (a) All eigenvectors of $A$

The eigenvalues of $A$ are roots of the characteristic polynomial $p_A(\lambda)$, hence the eigenvalues of $A$ are

$$\lambda = 1, -2$$

with algebraic multiplicities $1$ and $2$, respectively.

Let us first find the eigenvectors corresponding to the eigenvalue $\lambda = -2$.

The eigenvectors are nonzero solutions of the equation

$$(A + 2I)\mathbf{x} = 0,$$

where $I$ is the $3 \times 3$ identity matrix, $\mathbf{x} \in \mathbb{R}^3$, and $0$ is the three dimensional zero vector.

To solve the linear system, we use the Gauss-Jordan elimination method.

Applying elementary row operations we obtain

$$A + 2I = \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \rightarrow \cdots \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, the vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ must satisfy

$$x_1 = -x_2 - x_3.$$

Hence the eigenvectors associated to $\lambda = -2$ are

$$\mathrm{x} = \begin{bmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

for any pair of complex numbers $(x_1, x_2) \neq (0, 0)$.

Now we find the eigenvectors corresponding to the eigenvalue $\lambda = 1$.

As above, we solve the system

$$(A - I)\mathrm{x} = 0.$$

By elementary row operations, we reduce the matrix $A - I$ and obtain

$$A - I = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence the vector $\mathbf{x}$ should satisfy

$$x_1 = x_3 \text{ and } x_2 = -x_3.$$

Thus the eigenvectors associated to $\lambda = 1$ are

$$\mathrm{x} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

for any nonzero complex number $x_3$.

### (b) All eigenvectors of $B$

The eigenvalues of $B$ are roots of the characteristic polynomial $p_B(t)$, hence the eigenvalues of $B$ are

$$\lambda = 1, -2$$

with algebraic multiplicities $1$ and $2$, respectively.

Let us first find the eigenvectors corresponding to the eigenvalue $\lambda = -2$.

We need to solve the sysetm

$$(B + 2I)\mathrm{x} = 0.$$

By elementary row operations, we have

$$B + 2I = \begin{bmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus the vector $\mathbf{x}$ satisfies

$$x_1 = -x_2 \text{ and } x_3 = 0,$$

hence the eigenvectors associated to $\lambda = -2$ are

$$\mathrm{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

for any nonzero complex number $x_2$

---

Next, we find the eigenvectors corresponding to eigenvalue $\lambda = 1$.

Since we have

$$B - I = \begin{bmatrix} 1 & 4 & 3 \\ -4 & -7 & -3 \\ 3 & 3 & 0 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

by elementary row operations, the system $(B - I)\mathrm{x} = 0$ has solutions

$$x_1 = x_3 \text{ and } x_2 = -x_3.$$

Thus the eigenvectors associated to $\lambda = 1$ are

$$\mathrm{x} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

for any nonzero complex number $x_3$.

### (c) Which matrix $A$ or $B$ is diagonalizable?

From the result of (a), for each eigenvalue $\lambda$ of $A$ the geometric multiplicity of $\lambda$ is the same as the algebraic multiplicity of $\lambda$. Thus $A$ is not defective, and hence $A$ is diagonalizable.

From the result of (b), the geometric multiplicity of the eigenvalue $\lambda = -2$ of $B$ is $1$ although the algebraic multiplicity of $\lambda = -2$ is $2$. Thus the matrix $B$ is defective and hence $B$ is not diagonalizable.

### (d) Diagonalize the matrix

We diagonalize the matrix $A$. From the computation of (a), the vectors

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

form a basis of the eigenspace $E_{-2}$ associated to $\lambda = -2$.

Also the vector

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

form a basis of the eigenspace $E_1$ associated to $\lambda = 1$.

Therefore the vectors

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

are linearly independent eigenvectors of $A$.

Thus we set

$$P = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and we have

$$P^{-1}AP = D.$$

**32** Consider the matrix $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where $a$ and $b$ are real numbers and $b \neq 0$.

**(a)** Find all eigenvalues of $A$.

**(b)** For each eigenvalue of $A$, determine the eigenspace $E_\lambda$.

**(c)** Diagonalize the matrix $A$ by finding a nonsingular matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.

## Solution.

### (a) Find all eigenvalues of $A$.

The characteristic polynomial $p(t)$ of the matrix $A$ is

$$\begin{aligned} p(t) = \det(A - tI) &= \begin{vmatrix} a-t & -b \\ & a-t \end{vmatrix} \\ &= (a-t)^2 + b^2. \end{aligned}$$

The eigenvalues of $A$ are roots of $p(t)$.

So we solve $p(t) = 0$. We have

$$\begin{aligned} & (a-t)^2 + b^2 = 0 \\ \Leftrightarrow \quad & (a-t)^2 = -b^2 \\ \Leftrightarrow \quad & a - t = \pm ib \\ \Leftrightarrow \quad & t = a \pm ib. \end{aligned}$$

Here $i = \sqrt{-1}$.

Thus, the eigenvalues of $A$ are $a \pm ib$.

### (b) For each eigenvalue of $A$, determine the eigenspace $E_\lambda$.

We first determine the eigenspace $E_\lambda$ for $\lambda = a + ib$.
Recall that by definition $E_\lambda = \mathcal{N}(A - \lambda I)$, the nullspace of $A - \lambda I$.

We compute

$$A - (a+ib)I = \begin{bmatrix} -ib & -b \\ b & -ib \end{bmatrix} \xrightarrow{\frac{i}{b}R_1} \begin{bmatrix} 1 & -i \\ b & -ib \end{bmatrix} \xrightarrow{R_2 - bR_1} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}.$$

Note that in the above row reduction, we needed the assumption $b \neq 0$.

It follows that the general solution of the system is $x_1 = ix_2$.
Hence, we have

$$E_{a+ib} = \text{Span}\left( \begin{bmatrix} i \\ 1 \end{bmatrix} \right).$$

---

Note that the other eigenvalue $a - ib$ is the complex conjugate of $a + ib$.
It follows that the eigenspace $E_{a-ib}$ is obtained by conjugating the eigenspace $E_{a+ib}$.
Thus,

$$E_{a-ib} = \text{Span}\left( \begin{bmatrix} -i \\ 1 \end{bmatrix} \right).$$

### (c) Diagonalize the matrix $A$

From part (b), we see that

$$\begin{bmatrix} i \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

form an eigenbasis for $\mathbb{C}^2$.

So, we set

$$S = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} a+ib & 0 \\ 0 & a-ib \end{bmatrix},$$

and we obtain $S^{-1}AS = D$ by the diagonalization procedure.

33 Determine all $2 \times 2$ matrices $A$ such that $A$ has eigenvalues $2$ and $-1$ with corresponding eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, respectively.

respectively.

### Solution.

Suppose that $A$ is a $2 \times 2$ matrix having eigenvalues $2$ and $-1$ with corresponding eigenvectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

respectively.

Then since $A$ has two distinct eigenvalues, the matrix $A$ is diagonalizable.

As we know eigenvectors, we can diagonalize $A$ by the matrix

$$S := \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

That is, we have

$$S^{-1}AS = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}.$$

The inverse matrix of $S$ is given by

$$S^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

It follows that we have

$$\begin{aligned} A &= S \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} S^{-1} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -6 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Therefore, the only matrix satisfying the given conditions is

$$A = \begin{bmatrix} 2 & -6 \\ 0 & 1 \end{bmatrix}.$$

(34) Let $A$ and $B$ be $n \times n$ matrices. Suppose that $A$ and $B$ have the same eigenvalues $\lambda_1, \dots, \lambda_n$ with the same corresponding eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$. Prove that if the eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are linearly independent, then $A = B$.

**Proof.**

Since $A$ and $B$ have $n$ linearly independent eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, they are diagonalizable.

Specifically, if we put $S = [\mathbf{x}_1, \dots, \mathbf{x}_n]$.

Then $S$ is invertible (as column vectors of $S$ are linearly independent) and we have

$$S^{-1}AS = D \text{ and } S^{-1}BS = D,$$

where $D$ is the diagonal matrix whose diagonal entries are eigenvalues:

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

It follows that we have

$$S^{-1}AS = D = S^{-1}BS,$$

and hence $A = B$. This completes the proof.

---

㉟ Suppose that $A$ is a diagonalizable $n \times n$ matrix and has only $1$ and $-1$ as eigenvalues. Show that $A^2 = I_n$, where $I_n$ is the $n \times n$ identity matrix.
(*Stanford University*)

### Hint.

Diagonalize the matrix $A$ so that the diagonal matrix has $\pm 1$ on diagonal entries.

### Proof.

Since $A$ is diagonalizable, there exists an invertible matrix $P$ such that $P^{-1}AP = D$, where $D$ is a diagonal matrix.
Since $A$ has only $\pm 1$ as eigenvalues, we can choose $P$ so that the diagonal entries of $D$ are either $\pm 1$.

Then we have $A = PDP^{-1}$ and

$$A^2 = (PDP^{-1})(PDP^{-1}) = (PD\cancel{P}^{-1})(\cancel{P}DP^{-1}) = PD^2P^{-1}.$$

Note that $D^2 = I_2$ and thus we have $A^2 = I_2$ as required.

### Generalization

As a generalization, consider the following problem.

**Problem**.
Suppose that $A$ is a diagonalizable $n \times n$ matrix and the eigenvalues of $A$ are $r$-th roots of unity for some positive integer $r$.
Show that $A^r = I_n$, where $I_n$ is the $n \times n$ identity matrix.

The proof of Problem 2 is a straightforward generalization of the proof of Problem 1. So give it a shot.

---

㊱ Let

$$A = \begin{bmatrix} \frac{1}{7} & \frac{3}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{3}{7} & \frac{1}{7} \end{bmatrix}$$

be $3 \times 3$ matrix. Find

$$\lim_{n \to \infty} A^n.$$

(*Nagoya University Linear Algebra Exam*)

## Hint.

㊲ The matrix $A$ is symmetric, hence diagonalizable.

㊳ Diagonalize $A$.

㊴ You may want to find eigenvalues
without computing the characteristic polynomial for $A$.

## Solution.

Note that the matrix $A$ is symmetric, hence it is diagonalizable.

We observe several things to simplify the computation.

First, note that the sum of the entries in each row is $1$.

(Such a matrix is call a stochastic matrix. See the comment below.)

Thus the matrix $A$ has eigenvalue $1$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector.

Also note that if we add $2/7(=-\lambda)$ to diagonal entries, then every entry becomes $3/7$.

That is,

$$A+\frac{2}{7}I=\begin{bmatrix} \frac{3}{7} & \frac{3}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{3}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{3}{7} & \frac{3}{7} \end{bmatrix},$$

which is clearly singular.

From this observation, we notice that $-2/7$ is an eigenvalue of $A$ and eigenvectors are nonzero solution of $x_1+x_2+x_3=0$. Therefore $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ are linearly independent eigenvectors for the eigenvalue $-2/7$.

Hence the geometric (and hence algebraic) multiplicity of the eigenvalue $-2/7$ is $2$.

(Note that we found all eigenvalues of $A$ without actually computing the characteristic polynomial.)

Now the invertible matix $P=\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$ diagonalize $A$. Namely we have

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2/7 & 0 \\ 0 & 0 & -2/7 \end{bmatrix}.$$

Hence

$$A = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2/7 & 0 \\ 0 & 0 & -2/7 \end{bmatrix} P^{-1}.$$

Therefore we have

$$A^n = P \begin{bmatrix} 1^n & 0 & 0 \\ 0 & (-2/7)^n & 0 \\ 0 & 0 & (-2/7)^n \end{bmatrix} P^{-1} \text{ and } \lim_{n\to\infty} A^n = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^{-1}.$$

Find the inverse $P^{-1}$ by your favorite method

$$P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-1}{3} & \frac{-1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{-2}{3} \end{bmatrix}.$$

Then the answer is

$$\lim_{n\to\infty} A^n = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

## Comment.

You might come up with the idea to use the diagonalization as it is an exam problem in linear algebra.
But you soon realize that computing the characteristic polynomial for $A$ is a bit messy.
So sometimes it is important to find eigenvalues without finding the roots of the characteristic polynomial.

We observed that

(40) the sum of row entries of $A$ is $1$

(41) if we add $(2/7)I$ to $A$ then we obtain a matrix whose entries are all identical.

These observation yields the eigenvalues of $A$ without using the characteristic polynomial.

---

If the sum of entries of each row of a matrix is $1$, then the matrix is called called a **stochastic matrix** (or **Markov matrix**, **probability matrix**).
The matrix $A$ in the problem is an example of a stochastic matrix.

Stochastic matrices have always $1$ as an eigenvalue.

see the post

Stochastic matrix (Markov matrix) and its eigenvalues and eigenvectors

for an example of a $2 \times 2$ stochastic matrix.

42 Let

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

**(a)** Find the characteristic polynomial and all the eigenvalues (real and complex) of $A$. Is $A$ diagonalizable over the complex numbers?
**(b)** Calculate $A^{2009}$.
(*Princeton University*)

## Solution.

### (a) The characteristic polynomial and the eigenvalues

The characteristic polynomial $p(t)$ of the matrix $A$ is the determinant of $A - tI$. We compute $p(t) = \det(A - tI)$ as follows.

$$\begin{aligned} p(t) &= \det(A - tI) \\ &= \begin{vmatrix} -t & 0 & 1 \\ 1 & -t & 0 \\ 0 & 1 & -t \end{vmatrix} \\ &= -t \begin{vmatrix} -t & 0 \\ 1 & -t \end{vmatrix} + \begin{vmatrix} 1 & -t \\ 0 & 1 \end{vmatrix} \text{ by the first row cofactor expansion} \end{aligned}$$
$$= -t^3 + 1.$$

Thus, the characteristic polynomial of the matrix $A$ is

$$p(t) = -t^3 + 1.$$

The eigenvalues of the matrix $A$ is roots of the characteristic polynomial.

Hence solving $-t^3 + 1 = 0$, we obtain

$$t = 1, \frac{-1 \pm \sqrt{3}i}{2}$$

and these are all eigenvalues of $A$.

The matrix $A$ is a $3 \times 3$ matrix, and hence it has at most three distinct eigenvalues, and we found three distinct eigenvalues. In general, if an $n \times n$ matrix has $n$ distinct eigenvalues, the matrix is diagonalizable. Thus the matrix $A$ is diagonalizable.

### (b) $A^{2009}$

By direct computations, we have

$$A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$A^3 = AA^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = I,$$

where $I$ is the $3 \times 3$ identity matrix.

Therefore noting that $2009 = 3 \cdot 669 + 2$, we have

$$\begin{aligned} A^{2009} &= A^{3 \cdot 669 + 2} = (A^3)^{669} A^2 \\ &= I^{669} A^2 = A^2. \end{aligned}$$

Therefore, we obtain

$$A^{2009} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

## Another solution using diagonalization

Here, I give another solution for (b) using the diagonalization of the matrix $A$.
For this particular matrix $A$, the above solution is easier since the power of $A$ has a simple pattern.

The following computation is lengthy, but I give it for a pedagogical reason.
I just outline the solution and omit the detail computations.
We obtained eigenvectors $\zeta^k, k = 0, 1, 2$, where $\zeta$ is a primitive third root of unity in (a).
Eigenvectors are $\begin{bmatrix} \zeta^{2k} \\ \zeta^k \\ 1 \end{bmatrix}$
for each eigenvalue $\zeta^k$.
Then let

$$S = \begin{bmatrix} 1 & \zeta^2 & \zeta \\ 1 & \zeta & \zeta^2 \\ 1 & 1 & 1 \end{bmatrix}$$

be the matrix whose columns are eigenvectors.

It has the inverse

$$S^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ \zeta & \zeta^2 & 1 \\ \zeta^2 & \zeta & 1 \end{bmatrix}.$$

Then the matrix $S$ diagonalize the matrix $A$ and we obtain

$$S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{bmatrix}.$$

Then we compute

$$\begin{aligned} A^{2009} &= S \begin{bmatrix} 1 & 0 & 0 \\ 0 & \zeta^{2009} & 0 \\ 0 & 0 & \zeta^{2(2009)} \end{bmatrix} S^{-1} \\ &= S \begin{bmatrix} 1 & 0 & 0 \\ 0 & \zeta^{2} & 0 \\ 0 & 0 & \zeta \end{bmatrix} S^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \end{aligned}$$

㊸ Let $A$ be an $n \times n$ matrix with real number entries. Show that if $A$ is diagonalizable by an orthogonal matrix, then $A$ is a symmetric matrix.

## Proof.

Suppose that the matrix $A$ is diagonalizable by an orthogonal matrix $Q$.

The orthogonality of the matrix $Q$ means that we have

$$Q^{\mathrm{T}}Q = QQ^{\mathrm{T}} = I, \tag{*}$$

where $Q^{\mathrm{T}}$ is the transpose matrix of $Q$ and $I$ is the $n \times n$ identity matrix.

Since $Q$ diagonalizes the matrix $A$, we have

$$Q^{-1}AQ = D,$$

where $D$ is a diagonal matrix.

Equivalently, we have

$$A = QDQ^{-1}. \tag{**}$$

Taking transpose of both sides, we obtain

$$\begin{aligned} A^{\mathrm{T}} &= (QDQ^{-1})^{\mathrm{T}} \\ &= (Q^{-1})^{\mathrm{T}} D^{\mathrm{T}} Q^{-1} \\ &= (Q^{-1})^{\mathrm{T}} D Q^{-1} \text{ since } D \text{ is diagonal.} \end{aligned} \tag{***}$$

By (*), we observe that the inverse matrix of $Q$ is the transpose $Q^{\mathrm{T}}$, that is, $Q^{-1} = Q^{\mathrm{T}}$.

It follows from this observation and (***) that we have

$$A^{\mathrm{T}} = QDQ^{-1}.$$

(Note that $(Q^{-1})^{\mathrm{T}} = Q^{\mathrm{TT}} = Q$.)

Comparing this with (**), we obtain

$$A^{\mathrm{T}} = A,$$

and hence $A$ is a symmetric matrix.

44 Let

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Determine whether the matrix $A$ is diagonalizable. If it is diagonalizable, then diagonalize $A$.

## Solution 1.

We claim that the matrix $A$ is diagonalizable.

One way to see this is to note that $A$ is a real symmetric matrix, and hence it is diagonalizable.

Alternatively, we can compute eigenspaces and check whether $A$ is not defective (namely, the algebraic multiplicity and the geometric multiplicity of each eigenvalue of $A$ are the same.

---

To diagonalize the matrix $A$, we need to find eigenvalues and three linearly independent eigenvectors.

We compute the characteristic polynomial of $A$ as follows:

$$\begin{aligned} p(t) = &\det(A - tI) \\ = &\begin{bmatrix} 2-t & -1 & -1 \\ -1 & 2-t & -1 \\ -1 & -1 & 2-t \end{bmatrix} \\ = &(2-t)\begin{bmatrix} 2-t & -1 \\ -1 & 2-t \end{bmatrix} - (-1)\begin{bmatrix} -1 & -1 \\ -1 & 2-t \end{bmatrix} + (-1)\begin{bmatrix} -1 & 2-t \\ -1 & -1 \end{bmatrix} \\ &\text{(by the first row cofactor expansion)} \\ = &-t(t-3)^2. \end{aligned}$$

Since eigenvalues are the roots of the characteristic polynomial, eigenvalues of $A$ are 0 and 3 with algebraic multiplicity 1 and 2, respectively.

(If you did not confirm that $A$ is diagonalizable yet, then at this point we know that the geometric multiplicity of $\lambda = 0$ is 1 since the geometric multiplicity is alway greater than 0 and less than or equal to the algebraic multiplicity. However the geometric multiplicity of $\lambda = 3$ is either 1 or 2.)

---

Next, we determine the eigenspace $E_\lambda$ and its basis for each eigenvalue of $A$.

For $\lambda = 3$, we find solutions of $(A - 3I)\mathbf{x} = \mathbf{0}$.

We have

$$A - 3I = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \xrightarrow{\substack{R_2 - R_1 \\ R_3 - R_1}} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence a solution must satisfy $x_1 = -x_2 - x_3$, and thus the eigenspace is

$$E_3 = \left\{ \mathbf{x} \in \mathrm{R}^3 \;\middle|\; \mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ for any } x_2, x_3 \in \mathrm{C} \right\}.$$

From this expression, it is straightforward to check that the set

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis of $E_3$.

(Hence the geometric multiplicity of $\lambda = 3$ is 2, and $A$ is not defective and diagonalizable.)

---

For $\lambda = 0$, we solve $(A - 0I)\mathbf{x} = \mathbf{0}$, thus $A\mathbf{x} = \mathbf{0}$.

We have

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} -1 & 2 & -1 \\ 2 & -1 & -1 \\ -1 & -1 & 2 \end{bmatrix} \xrightarrow{-R_1} \begin{bmatrix} 1 & -2 & 1 \\ 2 & -1 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

$$\xrightarrow{\substack{R_2 - 2R_1 \\ R_3 + R_1}} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 3 & -3 \\ 0 & -3 & 3 \end{bmatrix} \xrightarrow{R_3 + R_2} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{3}R_2} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_1 + 2R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence any solution satisfies

$$x_1 = x_3 \text{ and } x_2 = x_3.$$

Therefore, the eigenspace is

$$E_0 = \left\{ \mathbf{x} \in \mathrm{R}^3 \;\middle|\; x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ for any } x_3 \in \mathrm{C} \right\},$$

and a basis of $E_0$ is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

---

Let

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Note that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis of $E_3$ and $\{\mathbf{u}_3\}$ is a basis of $E_0$.
Thus, it follows that the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent eigenvectors.
Put

$$S = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Since the columns of $S$ are linearly independent, the matrix $S$ is nonsingular. Then the procedure of the diagonalization yields that

$$S^{-1}AS = \begin{bmatrix} \mathbf{3} & 0 & 0 \\ 0 & \mathbf{3} & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix},$$

where diagonal entries are eigenvalues corresponding to the eigenvector $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ with this order.

## Solution 2.

The second solution uses a different method to find eigenvalues and eigenvectors.
Let $B = A - 3I$. Then every entry of $B$ is $-1$.
We find eigenvalues $\lambda$ and eigenvectors of $B$. Then the eigenvalues of $A$ are $\lambda + 3$ and eigenvectors are the same.

First we reduce the matrix $B$ as follows:

$$B = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \xrightarrow{\substack{R_2 - R_1 \\ R_3 - R_1}} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence the rank of $B$ is 1, and the nullity of $B$ is 2 by the rank-nullity theorem. It follows that $\lambda = 0$ is an eigenvalue of $B$.
Note that the eigenspace $E_0$ corresponding to $\lambda = 0$ is the null space of $A$. From the reduction above, we see that the null space consists of vectors

$$\mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

for any complex numbers $x_2, x_3$.
It follows that

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

are basis vectors of eigenspace $E_0$. Hence the geometric multiplicity of $\lambda = 0$ is 2.
(The algebraic multiplicity is either 2 or 3. We will see it must be 2.)

---

Since all the entries of $B$ are $-1$, by inspection, we find that the vector

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $-3$.
In fact, we have

$$B\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \\ -3 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Since the algebraic multiplicity of $\lambda = 0$ is either 2 or 3, and the sum of all the algebraic multiplicities is equal to 3, the algebraic multiplicity of $\lambda = -3$ must be 1 and that of $\lambda = 0$ is 2.
Hence the geometric multiplicity of $\lambda = -3$ is 1. Thus

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

is a basis vector of $E_{-3}$.

---

In a nutshell, we have obtained that eigenvalues of $B$ are 0 and $-3$ and basis vectors of $E_0$ and $E_{-3}$ are

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

respectively.

Since $A = B + 3I$, the eigenvalues of $A$ are $0 + 3 = 3$ and $-3 + 3 = 0$ and the corresponding eigenvectors are the same. Thus

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

are the basis vectors of the eigenspace $E_3$ and $E_0$ of $A$, respectively.

As in solution 1, we put

$$S = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Then we have

$$S^{-1}AS = \begin{bmatrix} \mathbf{3} & 0 & 0 \\ 0 & \mathbf{3} & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix},$$

where diagonal entries are eigenvalues corresponding to the eigenvector $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ with this order.

---

(45) Let $A$ be an $n \times n$ matrix with the characteristic polynomial

$$p(t) = t^3(t-1)^2(t-2)^5(t+2)^4.$$

Assume that the matrix $A$ is diagonalizable.
**(a)** Find the size of the matrix $A$.
**(b)** Find the dimension of the eigenspace $E_2$ corresponding to the eigenvalue $\lambda = 2$.
**(c)** Find the nullity of $A$.
(*The Ohio State University*)

## Hint/Definition.

(46) Recall that when a matrix is diagonalizable, the algebraic multiplicity of each eigenvalue is the same as the geometric multiplicity.

(47) The geometric multiplicity of an eigenvalue $\lambda$ is the dimension of the eigenspace $E_\lambda = \mathcal{N}(A - \lambda I)$ corresponding to $\lambda$.

(48) The nullity of $A$ is the dimension of the null space $\mathcal{N}(A)$ of $A$.

## Solution.

### (a) Find the size of the matrix $A$.

In general, if $A$ is an $n \times n$ matrix, then its characteristic polynomials has degree $n$.
Since the degree of $p(t)$ is $14$, the size of $A$ is $14 \times 14$.

### (b) Find the dimension of the eigenspace $E_2$ corresponding to the eigenvalue $\lambda = 2$.

Note that the dimension of the eigenspace $E_2$ is the geometric multiplicity of the eigenvalue $\lambda = 2$ by definition.

From the characteristic polynomial $p(t)$, we see that $\lambda = 2$ is an eigenvalue of $A$ with algebraic multiplicity $5$.
Since $A$ is diagonalizable, the algebraic multiplicity of each eigenvalue is the same as the geometric multiplicity.

It follows that the geometric multiplicity of $\lambda = 2$ is $5$, hence the dimension of the eigenspace $E_2$ is $5$.

(c) Find the nullity of $A$.

We first observe that $\lambda = 0$ is an eigenvalue of $A$ with algebraic multiplicity $3$ from the characteristic polynomial.

By definition, the nullity of $A$ is the dimension of the null space $\mathcal{N}(A)$, and furthermore the null space $\mathcal{N}(A)$ is the eigenspace $E_0$.
Thus, the nullity of $A$ is the same as the geometric multiplicity of the eigenvalue $\lambda = 0$.

Since $A$ is diagonalizable, the algebraic and geometric multiplicities are the same. Hence the nullity of $A$ is $3$.

---

(49) Let $A$ be an $n \times n$ real symmetric matrix whose eigenvalues are all non-negative real numbers. Show that there is an $n \times n$ real matrix $B$ such that $B^2 = A$.

---

## Hint.

(50) (a) Consider the eigenvalues and the Jordan canonical form of $A$

(51) (b) Diagonalize the matrix $A$

## Solution.

(a) Does there exist a $2 \times 2$ matrix $A$ with $A^3 = O$ but $A^2 \neq O$?

We claim that there is no such $2 \times 2$ matrix $A$.
Suppose that we have $A^3 = O$. This implies that all the eigenvalues of $A$ are zero.
To see this, let $\lambda$ be an eigenvalue of $A$ and let $\mathbf{x}$ be a corresponding eigenvector. Then applying the defining relation $A\mathbf{x} = \lambda\mathbf{x}$ successively, we have

$$0 = A^3\mathbf{x} = \lambda A^2\mathbf{x} = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}.$$

Thus, the eigenvalue $\lambda$ must be zero since $\mathbf{x}$ is a nonzero vector since it is an eigenvector.
Therefore, there exists an invertible matrix $P$ such that the Jordan canonical form of $A$ is

$P^{-1}AP = \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}$ for some number $a$.

Then $A = P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}P^{-1}$ and we have

$$\begin{aligned} A^2 &= \left(P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}P^{-1}\right)\left(P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}P^{-1}\right) = P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}^2 P^{-1} \\ &= P\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}P^{-1} = O. \end{aligned}$$

Hence $A^2$ must be zero, and we conclude that there is no $2 \times 2$ matrix $A$ satisfying both $A^3 = O$ and $A^2 \neq O$.

> (b) Does there exist a $3 \times 3$ real matrix $B$ such that $B^2 = A$?

The answer is yes and we will construct such a matrix $B$ as follows.

The characteristic polynomial $p(t)$ of the matrix $A$ is

$$p(t) := \det(A - tI) = -t(t-1)(t-3).$$

Thus the eigenvalues are $\lambda = 0, 1, 3$.
The corresponding eigenvalues are obtained by solving the system $(A - \lambda I)\mathrm{x} = 0$ and we see that

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

are eigenvectors corresponding to eigenvalues $0, 1, 3$, respectively.

Thus the matrix

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix}$$

is invertible and diagonalize the matrix $A$ such that

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Then if we have the equality $B^2 = A$, then we need to have

$$(P^{-1}BP)(P^{-1}BP) = p^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

From this we see that the matrix

$$B = P\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}P^{-1}$$

satisfies the above relation.

In fact, we compute

$$\begin{aligned} B^2 &= \left( P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} P^{-1} \right) \left( P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} P^{-1} \right) \\ &= P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}^2 P^{-1} = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} P^{-1} = A. \end{aligned}$$

Remark that the matrix $B$ is real since $P$ is real. Thus the matrix $B$ satisfies the conditions of the problem.

### Remark.

The explicit form of the matrix $P$ was used to assure that the matrix $B$ is real.
So we did not have to compute the matrix $P$ explicitly if we use the fact that the symmetric matrix is diagonalizable by a real orthogonal matrix.

### Generalization.

See problem A square root matrix of a symmetric matrix with non-negative eigenvalues for a more general question than part (b).

Furthermore, try the next problem.

(52) Find **a square root** of the matrix

$$A = \begin{bmatrix} 1 & 3 & -3 \\ 0 & 4 & 5 \\ 0 & 0 & 9 \end{bmatrix}.$$

How many square roots does this matrix have?
(*University of California, Berkeley*)

### Proof.

We will find all matrices $B$ such that $B^2 = A$. Such matrices $B$ are square roots of the matrix $A$.

Note that since $A$ is a diagonal matrix, the eigenvalues of $A$ are diagonal entries $1, 4, 9$. Since $A$ has three distinct eigenvalues, it is diagonalizable.
Solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$ for $\lambda = 1, 4, 9$, we find eigenvectors corresponding to eigenvalues $1, 4, 9$ are respectively

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Thus the invertible matrix

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

diagonalizes the matrix $A$, that is, we have

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

Then if $B^2 = A$, then we have $(P^{-1}BP)(P^{-1}B) = P^{-1}AP$.
Let $A' = P^{-1}AP$ and $B' = P^{-1}BP$.

Since we have $B'^2 = A'$, we have $B'A' = B'^3 = A'B'$.
Since $A'$ is diagonal with distinct diagonal entries, this implies that $B'$ is also a diagonal matrix.

A diagonal matrix $B'$ satisfying $B'^2 = A' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}$ is one of

$$\begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 2 & 0 \\ 0 & 0 & \pm 3 \end{bmatrix}.$$

Hence $B$ must be one of

$$P \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 2 & 0 \\ 0 & 0 & \pm 3 \end{bmatrix} P^{-1}.$$

The inverse matrix of $P$ can be calculated as

$$P^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore, all the square roots of the matrix $A$ are

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 2 & 0 \\ 0 & 0 & \pm 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

and we have $8$ square root matrices.

For example, when the diagonal matrix has all positive entries, then one of the square roots is

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

## Related Question.

**Problem**.

Prove that a positive definite matrix has a unique positive definite square root.

(53) Suppose the following information is known about a $3 \times 3$ matrix $A$.

$$A\begin{bmatrix}1\\2\\1\end{bmatrix}=6\begin{bmatrix}1\\2\\1\end{bmatrix}, \quad A\begin{bmatrix}1\\-1\\1\end{bmatrix}=3\begin{bmatrix}1\\-1\\1\end{bmatrix}, \quad A\begin{bmatrix}2\\-1\\0\end{bmatrix}=3\begin{bmatrix}1\\-1\\1\end{bmatrix}.$$

**(a)** Find the eigenvalues of $A$.
**(b)** Find the corresponding eigenspaces.
**(c)** Is $A$ a diagonalizable matrix? Is $A$ an invertible matrix? Is $A$ an idempotent matrix?
(*Johns Hopkins University*)

## Solution.

### (a) Find the eigenvalues of $A$.

From the first and the second equations, we see that $6$ and $3$ are eigenvalues of $A$.

From the second and the third equations, we have

$$A\begin{bmatrix}1\\-1\\1\end{bmatrix}=3\begin{bmatrix}1\\-1\\1\end{bmatrix}=A\begin{bmatrix}2\\-1\\0\end{bmatrix}$$

and thus we have

$$\begin{bmatrix}0\\0\\0\end{bmatrix}=A\left(\begin{bmatrix}1\\-1\\1\end{bmatrix}-\begin{bmatrix}2\\-1\\0\end{bmatrix}\right)=A\begin{bmatrix}-1\\0\\1\end{bmatrix}.$$

This yields that

$$A\begin{bmatrix}-1\\0\\1\end{bmatrix}=0\begin{bmatrix}-1\\0\\1\end{bmatrix} \tag{*}$$

and we conclude that $0$ is an eigenvalue.

Since the size of the matrix $A$ is $3 \times 3$, it has at most three eigenvalues. Thus we found them all.

Eigenvalues of $A$ are $0, 3, 6$.

### (b) Find the corresponding eigenspaces.

From the first equation, the vector $\begin{bmatrix}1\\2\\1\end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $6$.

We also see from the second equation, the vector $\begin{bmatrix}1\\-1\\1\end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $3$.

The equation (*) in the solution (a) also tells us that the vector $\begin{bmatrix}-1\\0\\1\end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $0$.

Note that we have three distinct eigenvalues for $3 \times 3$ matrix $A$. So all the algebraic multiplicities for eigenvalues are $1$, hence the geometric multiplicities must be $1$ since the geometric multiplicity is always less than or equal to the algebraic multiplicity.
Therefore, each eigenspace has dimension $1$. We already found a nonzero vector in each eigenspace, and thus the vector is a basis vector for each eigenspace.
We denote $E_\lambda$ for the eigenspace for eigenvector $\lambda$. Then we have

$$E_0 = \operatorname{Span}\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right), \quad E_3 = \operatorname{Span}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right), \quad E_6 = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right).$$

### (c) Diagonalizable matrix? Invertible matrix? Idempotent matrix?

Since $A$ has three distinct eigenvalues, $A$ is diagonalizable. (Or, algebraic multiplicities are the same as geometric multiplicities.)
Since the matrix $A$ has $0$ as an eigenvalue. Thus $A$ is not invertible.

If $A$ is idempotent matrix, then the eigenvalues of $A$ is either $0$ or $1$. Thus $A$ is not an idempotent matrix.
(For a proof of this fact, see the post Eigenvalues of an idempotent matrix.)

---

(54) Diagonalize the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Namely, find a nonsingular matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.
(*The Ohio State University*)

---

## Hint.

To diagonalize the matrix $A$, we need to find eigenvalues $A$ and bases of eigenspaces.

For a procedure of the diagonalization, see the post "How to Diagonalize a Matrix. Step by Step Explanation.".

Below, we will find eigenvalues and eigenvectors without using the characteristic polynomial although you may use it.

## Solution.

We use an indirect way to find eigenvalues and eigenvectors.
(We will not use the characteristic polynomial.)

Applying the elementary row operations, we have

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{R_2 - R_1 \\ R_3 - R_1}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence the solutions $\mathbf{x}$ of $A\mathbf{x} = 0$ satisfy

$$x_1 = -x_2 - x_3.$$

Thus, every vector in the null space is of the form

$$\mathbf{x} = \begin{bmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

for some scalars $x_2, x_3$.

It follows that

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis of the null space $\mathcal{N}(A)$ .

Hence $0$ is an eigenvalue and the geometric multiplicity corresponding to $0$, which is the nullity of $A$ , is $2$.
It follows that the algebraic multiplicity of the eigenvalue $0$ is either $2$ or $3$.
We see that it is $2$ shortly.

---

Note that by inspection we have

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

This yields that $3$ is an eigenvalue of $A$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is a corresponding eigenvector.

The sum of algebraic multiplicities of all eigenvalues of $A$ is $3$.
Hence the algebraic multiplicity of $0$ must be $2$, and that of $3$ must be $1$.
In particular, the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ forms a basis of the eigenspace $E_3$.

---

In summary so far, we have eigenvalues $0, 3$ and basis vectors of eigenspaces are

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

respectively.

Thus, we put

$$S = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

and obtain

$$S^{-1}AS = D,$$

where

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

(55) Prove that the matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is diagonalizable.
Prove, however, that $A$ cannot be diagonalized by a real nonsingular matrix.
That is, there is no real nonsingular matrix $S$ such that $S^{-1}AS$ is a diagonal matrix.

## Proof.

We first find the eigenvalues of $A$ by computing its characteristic polynomial $p(t)$.

We have

$$p(t) = \det(A - tI) = \begin{vmatrix} -t & 1 \\ -1 & -t \end{vmatrix} = t^2 + 1.$$

Solving $p(t) = t^2 + 1 = 0$, we obtain two distinct eigenvalues $\pm i$ of $A$.

Hence the matrix $A$ is diagonalizable.

---

To prove the second statement, assume, on the contrary, that $A$ is diagonalizable by a real nonsingular matrix $S$.

Then we have

$$S^{-1}AS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

by diagonalization.

As the matrices $A, S$ are real, the left-hand side is a real matrix.

Taking the complex conjugate of both sides, we obtain

$$\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = \overline{\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}} = \overline{S^{-1}AS} = S^{-1}AS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

This equality is clearly impossible.

Hence the matrix $A$ cannot be diagonalized by a real nonsingular matrix.

56 Let $A = \begin{bmatrix} 1-a & a \\ -a & 1+a \end{bmatrix}$ be a $2 \times 2$ matrix, where $a$ is a complex number. Determine the values of $a$ such that the matrix $A$ is diagonalizable.
(*Nagoya University*)

## Problem 459

(*Nagoya University, Linear Algebra Exam Problem*)

## Proof.

To find eigenvalues of the matrix $A$, we determine the characteristic polynomial $p(t)$ of $A$ as follows.
We have

$$\begin{aligned} p(t) = \det(A - tI) &= \begin{vmatrix} 1-a-t & a \\ -a & 1+a-t \end{vmatrix} \\ &= (1-a-t)(1+a-t) + a^2 \\ &= (1-t)^2 - a^2 + a^2 = (1-t)^2. \end{aligned}$$

Note
Thus, the eigenvalue of $A$ is $1$ with algebraic multiplicity $2$.

Let us determine the geometric multiplicity (the dimension of the eigenspace $E_1$).
We have

$$A - I = \begin{bmatrix} -a & a \\ -a & a \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} -a & a \\ 0 & 0 \end{bmatrix}.$$

If $a \neq 0$, then we further reduce it and get

$$\begin{bmatrix} -a & a \\ 0 & 0 \end{bmatrix} \xrightarrow{\frac{-1}{a} R_1} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

Hence the eigenspace corresponding to the eigenvalue $1$ is

$$E_1 = \mathcal{N}(A - I) = \mathrm{Span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$$

and its dimension is $1$, which is less than the algebraic multiplicity.
Thus, when $a \neq 0$, the matrix $A$ is not diagonalizable.

If $a = 0$, then the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is already diagonal, hence it is diagonalizable.

In conclusion, the matrix $A$ is diagonalizable if and only if $a = 0$.

57 Consider the $2 \times 2$ complex matrix $A = \begin{bmatrix} a & b-a \\ 0 & b \end{bmatrix}$.

**(a)** Find the eigenvalues of $A$.
**(b)** For each eigenvalue of $A$, determine the eigenvectors.
**(c)** Diagonalize the matrix $A$.
**(d)** Using the result of the diagonalization, compute and simplify $A^k$ for each positive integer $k$.

## Solution.

### (a) Find the eigenvalues of $A$.

Since $A$ is an upper triangular matrix, eigenvalues are diagonal entries.

Hence $a, b$ are eigenvalues of $A$.

### (b) For each eigenvalue of $A$, determine the eigenvectors.

Suppose now that $a \neq b$.

Let us find eigenvectors corresponding to the eigenvalue $a$.

We have

$$A - aI = \begin{bmatrix} 0 & b-a \\ 0 & b-a \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 0 & b-a \\ 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{b-a} R_1} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

It follows that the eigenvectors corresponding to $a$ are

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where $x_1$ is any nonzero complex number.

---

Next, we find the eigenvectors corresponding to the eigenvalue $b$.

We have

$$A - bI = \begin{bmatrix} a-b & b-a \\ 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{a-b} R_1} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

Hence the eigenvectors corresponding to $b$ are

$$x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $x_1$ is any nonzero complex number.

### (c) Diagonalize the matrix $A$.

When $a = b$, then $A$ is already diagonal matrix. So let us consider the case $a \neq b$.

In the previous parts, we obtained the eigenvalues $a, b$, and corresponding eigenvectors

$$\begin{bmatrix}1\\0\end{bmatrix} \text{ and } \begin{bmatrix}1\\1\end{bmatrix}.$$

Let $S = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$ be a matrix whose column vectors are the eigenvectors.

Then $S$ is invertible and we have

$$S^{-1}AS = \begin{bmatrix}a & 0\\0 & b\end{bmatrix}$$

by the diagonalization process.

Remark that this formula is also true even when $a = b$.

(d) Using the result of the diagonalization, compute and simplify $A^k$ for each positive integer $k$.

Using the result of the diagonalization in part (c), we have

$$A = S\begin{bmatrix}a & 0\\0 & b\end{bmatrix}S^{-1}.$$

For each positive integer $k$, we have

$$\begin{aligned}A^k &= \left(S\begin{bmatrix}a & 0\\0 & b\end{bmatrix}S^{-1}\right)^k\\ &= S\begin{bmatrix}a & 0\\0 & b\end{bmatrix}^k S^{-1} = S\begin{bmatrix}a^k & 0\\0 & b^k\end{bmatrix}S^{-1}\\ &= \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}a^k & 0\\0 & b^k\end{bmatrix}\begin{bmatrix}1 & -1\\0 & 1\end{bmatrix}\\ &= \begin{bmatrix}a^k & b^k - a^k\\0 & b^k\end{bmatrix}.\end{aligned}$$

In summary, we have the formula

$$A^k = \begin{bmatrix}a^k & b^k - a^k\\0 & b^k\end{bmatrix}.$$

58 Consider the complex matrix

$$A = \begin{bmatrix}\sqrt{2}\cos x & i\sin x & 0\\ i\sin x & 0 & -i\sin x\\ 0 & -i\sin x & -\sqrt{2}\cos x\end{bmatrix},$$

where $x$ is a real number between $0$ and $2\pi$. Determine for which values of $x$ the matrix $A$ is diagonalizable. When $A$ is diagonalizable, find a diagonal matrix $D$ so that $P^{-1}AP = D$ for some nonsingular matrix $P$.

## Solution.

Let us first find the eigenvalues of the matrix $A$.

To do so, we compute the characteristic polynomial $p(t) = \det(A - tI)$ of $A$ as follows.

Using Sarrus's rule to compute the $3 \times 3$ determinant, we have

$$\begin{aligned}
p(t) &= \det(A - tI) \\
&= \begin{bmatrix} \sqrt{2}\cos x - t & i\sin x & 0 \\ i\sin x & -t & -i\sin x \\ 0 & -i\sin x & -\sqrt{2}\cos x - t \end{bmatrix} \\
&= -t(\sqrt{2}\cos x - t)(-\sqrt{2}\cos x - t) - \big(-(\sin^2 x)(-\sqrt{2}\cos x - t) - (\sin^2 x)(\sqrt{2}\cos x - t)\big) \\
&= -t^3 + 2(\cos^2 x - \sin^2 x)t \\
&= -t^3 + 2\cos(2x)t.
\end{aligned}$$

The eigenvalues of $A$ are the roots of

$$p(t) = -t^3 + 2\cos(2x)t = -t(t^2 - 2\cos(2x)).$$

Hence the eigenvalues are

$$t = 0, \quad \pm\sqrt{2\cos(2x)}.$$

---

Note that if $\sqrt{2\cos(2x)} = -\sqrt{2\cos(2x)}$ then we have $\cos(2x) = 0$ and hence $x = \pi/4, 3\pi/4$ .
It follows that if $x = \pi/4, 3\pi/4$ , then the matrix $A$ has only one eigenvalue $0$ with algebraic multiplicity $3$.
Since $A$ is not the zero matrix, the rank of $A$ is greater than or equal to $1$.

Hence the nullity of $A$ is less than or equal to $2$ by the rank-nullity theorem.
It follows that the geometric multiplicity (=nullity) of the eigenvalue $0$ is strictly less than the algebraic multiplicity of $0$ and $A$ is not diagonalizable in this case.

---

Now suppose that $x \neq \pi/4, 3\pi/4$ .
In this case, the matrix $A$ has three distinct eigenvalues $0, \pm\sqrt{2\cos(2x)}$.
This implies that $A$ is diagonalizable.

Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ be eigenvectors corresponding to eigenvalues $0, \pm\sqrt{2\cos(2x)}$, respectively.
Define the $3 \times 3$ matrix $P$ by $P = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$.

It follows from the general procedure of the diagonalization that $P$ is a nonsingular matrix and

$$P^{-1}AP = D,$$

where $D$ is a diagonal matrix

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sqrt{2\cos(2x)} & 0 \\ 0 & 0 & -\sqrt{2\cos(2x)} \end{bmatrix}.$$

## Summary

In summary, when $x = \pi/4, 3\pi/4$ the matrix $A$ is not diagonalizable.

When $x \neq \pi/4, 3\pi/4$ , the matrix $A$ is diagonalizable and we can take the diagonal matrix $D$ as

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sqrt{2\cos(2x)} & 0 \\ 0 & 0 & -\sqrt{2\cos(2x)} \end{bmatrix}.$$

(59) Consider the Hermitian matrix $A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$.

**(a)** Find the eigenvalues of $A$.

**(b)** For each eigenvalue of $A$, find the eigenvectors.

**(c)** Diagonalize the Hermitian matrix $A$ by a unitary matrix. Namely, find a diagonal matrix $D$ and a unitary matrix $U$ such that $U^{-1}AU = D$.

---

## Solution.

### (a) Find the eigenvalues of $A$.

To find the eigenvalues of the matrix $A$, we first compute the characteristic polynomial $p(t)$ of $A$.

We have

$$\begin{aligned} p(t) = \det(A - tI) &= \begin{vmatrix} 1-t & i \\ -i & 1-t \end{vmatrix} \\ &= (1-t)(1-t) - (i)(-i) = t^2 - 2t + 1 - 1 \\ &= t(t-2). \end{aligned}$$

Solving $p(t) = 0$, we obtain the eigenvalues $0$ and $2$.

### (b) For each eigenvalue of $A$, find the eigenvectors.

Let us first find the eigenvectors corresponding to the eigenvalue $0$.

We solve the equation $(A - 0I)\mathrm{x} = 0$ .

We have

$$A \xrightarrow{R_2 + iR_1} \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}.$$

Hence the eigenvalues are

$$c\begin{bmatrix} 1 \\ i \end{bmatrix}$$

for any nonzero complex number $c$.

---

Next, we find the eigenvectors corresponding to the eigenvalue $2$. We solve the equation $(A - 2I)\mathrm{x} = 0$
.

We have

$$A - 2I = \begin{bmatrix} -1 & i \\ -i & -1 \end{bmatrix} \xrightarrow{-R_1}$$

$$\begin{bmatrix} 1 & -i \\ -i & -1 \end{bmatrix} \xrightarrow{R_2 + iR_1} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}.$$

Hence the eigenvalues are

$$c \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

for any nonzero complex number $c$.

### (c) Diagonalize the Hermitian matrix $A$ by a unitary matrix.

To diagonalize the Hermitian matrix $A$ by a unitary matrix $U$, we find an orthonormal basis for each eigenspace of $A$.
As each eigenspace of $A$ is $1$-dimensional by part (b), we just need to normalize any eigenvector for each eigenvalue.

By part (b), we know that $\mathbf{v}_1 := \begin{bmatrix} 1 \\ i \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $0$.
The length of this vector is

$$\|\mathbf{v}_1\| = \sqrt{1 \cdot 1 + (i)(-i)} = \sqrt{2}.$$

Hence the vector

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

is a unit eigenvector.

Similarly, from (b) we see that $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $2$. The length of the vector is $\sqrt{2}$.
Hence

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

is a unit eigenvector.

It follows that the matrix

$$U = [\mathbf{u}_1 \quad \mathbf{u}_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

is unitary and

$$U^{-1}AU = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

by diagonalization process.

(60) Let $A$ be an $n \times n$ complex matrix. Let $S$ be an invertible matrix.
**(a)** If $SAS^{-1} = \lambda A$ for some complex number $\lambda$, then prove that either $\lambda^n = 1$ or $A$ is a singular matrix.
**(b)** If $n$ is odd and $SAS^{-1} = -A$, then prove that $0$ is an eigenvalue of $A$.
**(c)** Suppose that all the eigenvalues of $A$ are integers and $\det(A) > 0$. If $n$ is odd and $SAS^{-1} = A^{-1}$, then prove that $1$ is an eigenvalue of $A$.

## Proof.

### Basic Properties of Determinants

We use the following properties of determinants of matrices.

For $n \times n$ matrices $A, B$ , we have

(61) $\det(AB) = \det(A)\det(B)$ .

(62) $\det(A^{-1}) = \det(A)^{-1}$ if $A$ is invertible.

(63) $\det(cA) = c^n \det(A)$ for any complex number $c$.

### (a) If $SAS^{-1} = \lambda A$ , then prove that $\lambda^n = 1$ or $A$ is a singular matrix.

Suppose that we have $SAS^{-1} = \lambda A$ .

We consider the determinants of both sides and we have

$$\det(SAS^{-1}) = \det(\lambda A).$$

The left hand side becomes

$$\begin{aligned} \det(SAS^{-1}) &= \det(S)\det(A)\det(S^{-1}) && \text{by property 1} \\ &= \det(S)\det(A)\det(S)^{-1} && \text{by property 2} \\ &= \det(A). \end{aligned}$$

The right hand side is by property 3

$$\det(\lambda A) = \lambda^n \det(A).$$

Hence we obtain

$$\det(A) = \lambda^n \det(A).$$

It follows from this that either $\lambda^n = 1$ or $\det(A) = 0$.

In conclusion, either $\lambda^n = 1$ or $A$ is a singular matrix.

(b) If $n$ is odd and $SAS^{-1} = -A$, then prove that $0$ is an eigenvalue of $A$.

Suppose that we have $SAS^{-1} = -A$.

Then we have

$$\begin{aligned} \det(A) &= \det(S)\det(A)\det(S)^{-1} \\ &= \det(S)\det(A)\det(S^{-1}) && \text{by property 2} \\ &= \det(SAS^{-1}) && \text{by property 1} \\ &= \det(-A) && \text{by assumption} \\ &= (-1)^n \det(A) && \text{by property 3} \\ &= -\det(A) && \text{since } n \text{ is odd.} \end{aligned}$$

This yields that $\det(A) = 0$.

Note that the product of all eigenvalues of $A$ is $\det(A)$.

(See the post "Determinant/Trace and Eigenvalues of a Matrix" for a proof.)

Thus, $0$ is an eigenvalue of $A$.

(c) Note that as $\det(A) > 0$, the matrix $A$ is invertible.
Suppose that $SAS^{-1} = A^{-1}$.
Then we have

$$\begin{aligned} \det(A) &= \det(S)\det(A)\det(S)^{-1} \\ &= \det(S)\det(A)\det(S^{-1}) && \text{by property 2} \\ &= \det(SAS^{-1}) && \text{by property 1} \\ &= \det(A^{-1}) && \text{by assumption} \\ &= \det(A)^{-1} && \text{by property 2.} \end{aligned}$$

Thus we have $\det(A)^2 = 1$, hence $\det(A) = 1$ as $\det(A) > 0$ by assumption.

---

Note again that the product of all eigenvalues of $A$ is $\det(A)$.
Since the eigenvalues of $A$ are integers by assumption, the eigenvalues of $A$ are either $1$ or $-1$.

If all of the $n$ eigenvalues of $A$ are $-1$, then the determinant is

$$\det(A) = (-1)^n = -1$$

since $n$ is odd, and this is a contradiction.
Thus, at least one of the eigenvalues must be $1$.

This completes the proof.

## Similar Transformation (conjugate)

A transformation $A$ to $SAS^{-1}$, for some invertible matrix $S$, is called a **similarity transformation** or **conjugation** of the matrix $A$.

We say that matrices $A$ and $B$ are **similar** if there exists an invertible matrix $S$ such that $B = SAS^{-1}$ . In other words, $A$ is similar to $B$ if there is a similarity transformation from $A$ to $B$ .

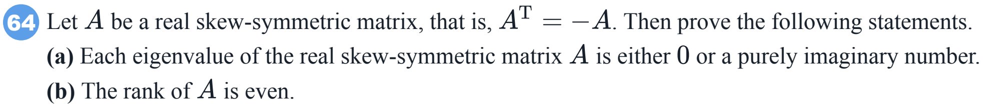
Check out the following problems about similar matrices.

**Problem**. Is the matrix $A = \begin{bmatrix} -1 & 6 \\ -2 & 6 \end{bmatrix}$ similar to the matrix $B = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$ ?

For a solution together with similar problems, see the post "Determine whether given matrices are similar".

**Problem**. Let $A, B$ be matrices. Show that if $A$ is diagonalizable and if $B$ is similar to $A$ , then $B$ is diagonalizable.

For a proof, see the post "A matrix similar to a diagonalizable matrix is also diagonalizable".



64 Let $A$ be a real skew-symmetric matrix, that is, $A^{\mathrm{T}} = -A$. Then prove the following statements.
**(a)** Each eigenvalue of the real skew-symmetric matrix $A$ is either $0$ or a purely imaginary number.
**(b)** The rank of $A$ is even.

## Proof.

### (a) Each eigenvalue of the real skew-symmetric matrix $A$ is either $0$ or a purely imaginary number.

Let $\lambda$ be an eigenvalue of $A$ and let $\mathbf{x}$ be an eigenvector corresponding to the eigenvalue $\lambda$ . That is, we have

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Multiplying by $\bar{\mathbf{x}}^{\mathrm{T}}$ from the left, we have

$$\bar{\mathbf{x}}^{\mathrm{T}} A\mathbf{x} = \lambda\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{x} = \lambda||\mathbf{x}||^2. \tag{*}$$

Note that the left hand side $\bar{\mathbf{x}}^{\mathrm{T}} A\mathbf{x}$ is the dot (inner) product of $\bar{\mathbf{x}}$ and $A\mathbf{x}$ . Since the dot product is commutative, we have

$$\begin{aligned} &\text{The left hand side of } (*) \\ &= \bar{\mathbf{x}}^{\mathrm{T}} A\mathbf{x} = (A\mathbf{x})^{\mathrm{T}}\bar{\mathbf{x}} \\ &= x^{\mathrm{T}} A^{\mathrm{T}}\bar{\mathbf{x}}. \end{aligned}$$

Since $A$ is skew-symmetric, we have $A^{\mathrm{T}} = -A$ . Substituting this into the above equality, we have

$$\begin{aligned} &\text{The left hand side of } (*) \\ &= x^{\mathrm{T}} A^{\mathrm{T}}\bar{\mathbf{x}} = -\mathbf{x}^{\mathrm{T}} A\bar{\mathbf{x}} \end{aligned}$$

---

Taking conjugate of $A\mathbf{x} = \lambda\mathbf{x}$ and use the fact that $A$ is real, we have

$$A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}.$$

Thus, we have

$$\begin{aligned}&\text{The left hand side of (*)}\\&=-\mathbf{x}^{\mathrm{T}}A\bar{\mathbf{x}}\\&=-\mathbf{x}^{\mathrm{T}}\bar{\lambda}\bar{\mathbf{x}}=-\bar{\lambda}||\mathbf{x}||^2.\end{aligned}$$

Therefore comparing the left and right hand sides of (*) yields

$$-\bar{\lambda}||\mathbf{x}||^2=\lambda||\mathbf{x}||^2.$$

Since $\mathbf{x}$ is an eigenvector, it is nonzero by definition. Thus $||\mathbf{x}||\neq 0$.

---

Hence we have

$$-\bar{\lambda}=\lambda,$$

and this implies that $\lambda$ is either $0$ or purely imaginary number.
(To see this, let $\lambda=a+ib$, where $a,b\in\mathbb{R}$. Then

$$-\bar{\lambda}=-a+ib=a+ib=\lambda$$

implies $a=0$, thus $\lambda=bi$.)

Remark: Another similar proof is to take conjugate of (*).

### (b) The rank of $A$ is even

From part (a), we know that the eigenvalues of $A$ are $0$ or purely imaginary.

Thus if $\lambda$ is a purely imaginary eigenvalue of $A$, then its conjugate $\bar{\lambda}=-\lambda$ is also an eigenvalue of $A$ since $A$ is a real matrix.
Thus, nonzero eigenvalues come in pairs $\lambda,-\lambda$ (and their algebraic multiplicities are the same).

Let

$$\lambda_1,-\lambda_1,\lambda_2,-\lambda_2,\dots,\lambda_k,-\lambda_k$$

be nonzero eigenvalues of $A$.

---

Since a real skew-symmetric matrix is normal, it is diagonalizable (by a unitary matrix).
Thus there exists an invertible matrix $P$ such that

$$\begin{aligned}&P^{-1}AP=\\&\operatorname{diag}[\lambda_1\quad -\lambda_1\quad \lambda_2\quad -\lambda_2\quad \dots\quad \lambda_k\quad -\lambda_k\quad 0\quad \dots 0],\end{aligned}$$

where $\operatorname{diag}[a_1,a_2,\dots,a_n]$ denotes the $n\times n$ matrix whose diagonal entries are $a_1,a_2,\dots,a_n$ and all the off diagonal entries are zero.

---

Since $P$ is an invertible matrix, the rank of $A$ is the same as the rank of the diagonal matrix on the right-hand side, which is easily seen to be $2k$.
Thus the rank of $A$ is $2k$, and we have proved that the rank of the real skew-symmetric matrix $A$ is even.

## Related Question.

As an application of this problem, try the following problem.

**Problem 1**. If $A$ is a skew-symmetric matrix, then $I+A$ and $I-A$ are nonsingular matrices, where $I$ is the identity matrix.

For a proof, see the post↴
If $A$ is a Skew-Symmetric Matrix, then $I+A$ is Nonsingular and $(I-A)(I+A)^{-1}$ is Orthogonal.

**Problem 2**. Prove that the determinant of an $n \times n$ skew-symmetric matrix is $0$ if $n$ is odd.

For a proof, check out the post ↴
The Determinant of a Skew-Symmetric Matrix is Zero.

## More Eigenvalue and Eigenvector Problems

Problems about eigenvalues and eigenvectors are collected on the page:

Eigenvectors and Eigenspaces

(65) Let $A$ be an $n \times n$ real symmetric matrix. Prove that there exists an eigenvalue $\lambda$ of $A$ such that for any vector $\mathbf{v} \in \mathbb{R}^n$, we have the inequality $\mathbf{v} \cdot A\mathbf{v} \leq \lambda \|\mathbf{v}\|^2$.

## Proof.

Recall that all the eigenvalues of a symmetric matrices are real numbers.
Let $\lambda_1, \dots, \lambda_n$ be eigenvalues of $A$.

Since these eigenvalues are real numbers, there is the largest one.
Let $\lambda$ be the largest eigenvalue of $A$.
With this choice of $\lambda$ we show that the inequality

$$\mathbf{v} \cdot A\mathbf{v} \leq \lambda \|\mathbf{v}\|^2$$

holds for any $\mathbf{v} \in \mathrm{R}^n$.

---

Also recall that for a real symmetric matrix, there are eigenvalues $\mathbf{v}_1, \dots, \mathbf{v}_n$ corresponding to $\lambda_1, \dots, \lambda_n$ such that

$$B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$$

form an orthonormal basis of $\mathrm{R}^n$.

(This statement is equivalent to that every real symmetric matrix is diagonalizable by an orthogonal matrix.)

---

Let $\mathbf{v}$ be an arbitrary vector in $\mathrm{R}^n$.

Then since $B$ is a basis of $\mathrm{R}^n$, we can write

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

for some $c_1, \ldots, c_n \in \mathrm{R}$.

Then we calculate

$$\begin{aligned} A\mathbf{v} &= A(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) \\ &= c_1A\mathbf{v}_1 + \cdots + c_nA\mathbf{v}_n \\ &= c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n \end{aligned}$$

since $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for $i = 1, \ldots, n$.

---

Using this, we have

$$\begin{aligned} \mathbf{v} \cdot A\mathbf{v} &= (c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) \cdot (c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n) \\ &= c_1^2\lambda_1 + \cdots + c_n^2\lambda_n. \end{aligned}$$

Here, we used that $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of $\mathrm{R}^3$.

That is, we used the properties

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

---

Since $\lambda$ is the largest eigenvalue of $A$, we have

$$\begin{aligned} \mathbf{v} \cdot A\mathbf{v} &= c_1^2\lambda_1 + \cdots + c_n^2\lambda_n \\ &\leq c_1^2\lambda + \cdots + c_n^2\lambda \\ &= \lambda(c_2^2 + \cdots + c_n^2) \\ &= \lambda\|\mathbf{v}\|^2. \end{aligned}$$

Hence the required inequality holds.

66 A real symmetric $n \times n$ matrix $A$ is called **positive definite** if $\mathbf{x}^{\mathrm{T}}A\mathbf{x} > 0$ for all nonzero vectors $\mathbf{x}$ in $\mathbb{R}^n$.

**(a)** Prove that the eigenvalues of a real symmetric positive-definite matrix $A$ are all positive.

**(b)** Prove that if eigenvalues of a real symmetric matrix $A$ are all positive, then $A$ is positive-definite.

### Proof.

(a) Prove that the eigenvalues of a real symmetric positive-definite matrix $A$ are all positive.

Recall that the eigenvalues of a real symmetric matrix are real.

(See the corollary in the post "Eigenvalues of a Hermitian matrix are real numbers".)

Let $\lambda$ be a (real) eigenvalue of $A$ and let $\mathbf{x}$ be a corresponding real eigenvector. That is, we have

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Then we multiply by $\mathbf{x}^{\mathrm{T}}$ on left and obtain

$$\begin{aligned} \mathbf{x}^{\mathrm{T}} A\mathbf{x} &= \lambda\mathbf{x}^{\mathrm{T}}\mathbf{x} \\ &= \lambda ||\mathbf{x}||^2. \end{aligned}$$

The left hand side is positive as $A$ is positive definite and $\mathbf{x}$ is a nonzero vector as it is an eigenvector.

Since the length $||\mathbf{x}||^2$ is positive, we must have $\lambda$ is positive.

It follows that every eigenvalue $\lambda$ of $A$ is real.

(b) Prove that if eigenvalues of a real symmetric matrix $A$ are all positive, then $A$ is positive-definite.

Note that a real symmetric matrix is diagonalizable by an orthogonal matrix.

So there exists an on orthogonal matrix $Q$ such that $Q^{\mathrm{T}} AQ = D$ , where $D$ is the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

whose diagonal entries $\lambda_i$ are eigenvalues of $A$ , which are positive by assumption.

Let $\mathbf{x}$ be an arbitrary nonzero vector in $\mathbb{R}^n$ .

Since $A = QDQ^{\mathrm{T}}$ (remark that $Q^{-1} = Q^{\mathrm{T}}$), we have

$$\mathbf{x}^{\mathrm{T}} A\mathbf{x} = \mathbf{x}^{\mathrm{T}} QDQ^{\mathrm{T}}\mathbf{x}.$$

Putting $\mathbf{y} = Q^{\mathrm{T}}\mathbf{x}$ , we can rewrite the above equation as

$$\mathbf{x}^{\mathrm{T}} A\mathbf{x} = \mathbf{y}^{\mathrm{T}} D\mathbf{y}.$$

Let

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Then we have

$$\begin{aligned}
\mathbf{x}^{\mathrm{T}} A \mathbf{x} &= \mathbf{y}^{\mathrm{T}} D \mathbf{y} \\
&= \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\
&= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 .
\end{aligned}$$

By assumption eigenvalues $\lambda_i$ are positive.
Also, since $\mathbf{x}$ is a nonzero vector and $Q$ is invertible, $\mathbf{y} = Q^{\mathrm{T}}\mathbf{x}$ is not a zero vector.
Thus the sum expression above is positive, hence $\mathbf{x}^{\mathrm{T}} A \mathbf{x}$ is positive for any nonzero vector $\mathbf{x}$.
Therefore, the matrix $A$ is positive-definite.

## Related Questions.

**Problem**. Let $A$ be an $n \times n$ real matrix. Prove the followings.

**(a)** The matrix $AA^{\mathrm{T}}$ is a symmetric matrix.

**(b)** The set of eigenvalues of $A$ and the set of eigenvalues of $A^{\mathrm{T}}$ are equal.

**(c)** The matrix $AA^{\mathrm{T}}$ is non-negative definite.
(An $n \times n$ matrix $B$ is called *non-negative definite* if for any $n$ dimensional vector $\mathbf{x}$, we have $\mathbf{x}^{\mathrm{T}} B \mathbf{x} \geq 0$.)

**(d)** All the eigenvalues of $AA^{\mathrm{T}}$ is non-negative.

For a solution, see the post "Transpose of a matrix and eigenvalues and related questions.".

**Problem**. Suppose $A$ is a positive definite symmetric $n \times n$ matrix.

**(a)** Prove that $A$ is invertible.

**(b)** Prove that $A^{-1}$ is symmetric.

**(c)** Prove that $A^{-1}$ is positive-definite.

For proofs, see the post "Inverse matrix of positive-definite symmetric matrix is positive-definite".

**Problem**.
Prove that a positive definite matrix has a unique positive definite square root.

For a solution of this problem, see the post
A Positive Definite Matrix Has a Unique Positive Definite Square Root

67 Suppose $A$ is a positive definite symmetric $n \times n$ matrix.
**(a)** Prove that $A$ is invertible.
**(b)** Prove that $A^{-1}$ is symmetric.
**(c)** Prove that $A^{-1}$ is positive-definite.
(*MIT*)

## Proof.

### (a) Prove that $A$ is invertible.

Recall that a symmetric matrix is positive-definite if and only if its eigenvalues are all positive.

Thus, since $A$ is positive-definite, the matrix does not have $0$ as an eigenvalue.
Hence $A$ is invertible.

### (b) Prove that $A^{-1}$ is symmetric.

By part (a), we know that $A$ is invertible. We have

$$A^{-1}A = I,$$

where $I$ is the $n \times n$ identity matrix.

Taking the transpose, we have

$$\begin{aligned} I = I^{\mathrm{T}} &= (A^{-1}A)^{\mathrm{T}} \\ &= A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} \\ &= A(A^{-1})^{\mathrm{T}} && \text{since } A \text{ is symmetric.} \end{aligned}$$

It follows that $A^{-1} = (A^{-1})^{\mathrm{T}}$, and hence $A^{-1}$ is a symmetric matrix.

### (c) Prove that $A^{-1}$ is positive-definite.

Again we use the fact that a symmetric matrix is positive-definite if and only if its eigenvalues are all positive.
(See the post "Positive definite real symmetric matrix and its eigenvalues" for a proof.)

All eigenvalues of $A^{-1}$ are of the form $1/\lambda$, where $\lambda$ is an eigenvalue of $A$.
Since $A$ is positive-definite, each eigenvalue $\lambda$ is positive, hence $1/\lambda$ is positive.

So all eigenvalues of $A^{-1}$ are positive, and it yields that $A^{-1}$ is positive-definite.

68 Let $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.
**(a)** Find eigenvalues of the matrix $A$.
**(b)** Find eigenvectors for each eigenvalue of $A$.
**(c)** Diagonalize the matrix $A$. That is, find an invertible matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$.
**(d)** Diagonalize the matrix $A^3 - 5A^2 + 3A + I$, where $I$ is the $2 \times 2$ identity matrix.
**(e)** Calculate $A^{100}$. (You do not have to compute $5^{100}$.)
**(f)** Calculate $(A^3 - 5A^2 + 3A + I)^{100}$. Let $w = 2^{100}$. Express the solution in terms of $w$.

## Solution.

### (a) Find eigenvalues of the matrix $A$ .

To find the eigenvalues of $A$ , we calculate the characteristic polynomial $p(t)$ as follows.
We have

$$\begin{aligned} p(t) = \det(A - tI) &= \begin{vmatrix} 1-t & 2 \\ 4 & 3-t \end{vmatrix} \\ &= (1-t)(3-t) - 8 = t^2 - 4t - 5 = (t+1)(t-5). \end{aligned}$$

The eigenvalues of $A$ are roots of its characteristic polynomial $p(t)$ .
Hence the eigenvalues of $A$ are $-1$ and $5$.

### (b) Find eigenvectors for each eigenvalue of $A$ .

We first determine the eigenvectors of the eigenvalue $-1$ by solving the system $(A+I)\mathrm{x} = 0$ .
We have

$$A + I = \begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix} \xrightarrow[\text{then } \frac{1}{2}R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

This yields that the eigenvectors corresponding to $-1$ are

$$a \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

for any nonzero scalar $a$.

Next, we find the eigenvectors corresponding to the eigenvalue $5$ by solving $(A - 5I)\mathrm{x} = 0$ .
We have

$$A - 5I = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix} \xrightarrow[\text{then } \frac{-1}{4}R_1]{R_2 + R_1} \begin{bmatrix} 1 & -1/2 \\ 0 & 0 \end{bmatrix}.$$

It follows that the eigenvectors corresponding to $5$ are

$$a \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

for any nonzero scalar $a$.

### (c) Diagonalize the matrix $A$ .

From part (a) and part (b), we have seen that $A$ has eigenvalues $-1$ and $5$ with corresponding eigenvectors

$$\mathrm{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \mathrm{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

(Here we chose the scalars $a$ to be $1$ but you could use any nonzero values for the scalars $a$.)

Let

$$S = [\mathbf{u} \quad \mathbf{v}] = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}.$$

Then the general procedure of the diagonalization yields that the matrix $S$ is invertible and

$$S^{-1}AS = D,$$

where $D$ is the diagonal matrix given by

$$D = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}.$$

(d) Diagonalize the matrix $A^3 - 5A^2 + 3A + I$.

In part (c), we obtained

$$S^{-1}AS = D,$$

where

$$S = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \text{ and } D = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}.$$

Note that we have $A = SDS^{-1}$ and

$$\begin{aligned} A^2 &= AA = SDS^{-1} \cdot SDS^{-1} = SD^2S^{-1} \\ A^3 &= A^2A = SD^2S^{-1} \cdot SDS^{-1} = SD^3S^{-1}. \end{aligned}$$

These relations gives

$$\begin{aligned} A^3 - 5A^2 + 3A + I &= SD^3S^{-1} - 5SD^2S^{-1} + 3SDS^{-1} + I \\ &= S(D^3 - 5D^2 + 3D + I)S^{-1}. \end{aligned}$$

Hence we obtain

$$\begin{aligned} &S^{-1}(A^3 - 5A^2 + 3A + I)S \\ &= D^3 - 5D^2 + 3D + I \\ &= \begin{bmatrix} (-1)^3 & 0 \\ 0 & 5^3 \end{bmatrix} - 5\begin{bmatrix} (-1)^2 & 0 \\ 0 & 5^2 \end{bmatrix} + 3\begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -8 & 0 \\ 0 & 16 \end{bmatrix}. \end{aligned}$$

# The Cayley-Hamilton Theorem

## Summary

1. (**The Cayley-Hamilton Theorem**) If $p(t)$ is the characteristic polynomial for an $n \times n$ matrix $A$, then the matrix $p(A)$ is the $n \times n$ zero matrix.

## Example

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$. The characteristic polynomial $p(t)$ of $A$ is

$$\begin{aligned} p(t) = \det(A - tI) &= \begin{bmatrix} 1-t & 1 \\ 1 & 3-t \end{bmatrix} \\ &= t^2 - 4t + 2. \end{aligned}$$

Then the Cayley-Hamilton theorem says that the matrix $p(A) = A^2 - 4A + 2I$ is the $2 \times 2$ zero matrix. In fact, we can directly check this:

$$\begin{aligned} p(A) = A^2 - 4A + 2I &= \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} - 4 \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} + 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 4 \\ 4 & 10 \end{bmatrix} + \begin{bmatrix} -4 & -4 \\ -4 & -12 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

## Problems

1. Let $T = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. Calculate and simplify the expression $-T^3 + 4T^2 + 5T - 2I$, where $I$ is the $3 \times 3$ identity matrix.

### Hint.

Use the ***Cayley-Hamilton theorem***.

> If $p(t)$ is the characteristic polynomial for an $n \times n$ matrix $A$, then the matrix $p(A)$ is the $n \times n$ zero matrix.

### Solution.

We use the Cayley-Hamilton theorem.

**The Cayley-Hamilton Theorem**.
If $p(t)$ is the characteristic polynomial for an $n \times n$ matrix $A$, then the matrix $p(A)$ is the $n \times n$ zero matrix.

To obtain the characteristic polynomial for $T$, we note that the matrix $T$ is upper triangular. Thus $T - tI$ is also upper triangular and recall that the determinant of an upper triangular matrix is the product of the diagonal entries. Thus the characteristic polynomial $p_T(t)$ for $T$ is

$$p_T(t) = \det(T - tI) = (1-t)(1-t)(2-t) = -t^3 + 4t^2 - 5t + 2.$$

---

By the Cayley-Hamilton theorem, we have

$$p_T(T) = -T^3 + 4T^2 - 5T + 2I = O.$$

(Don't forget $I$.)
Here $O$ is the $3 \times 3$ zero matrix.

---

Now we compute

$$\begin{aligned} -T^3 + 4T^2 + 5T - 2I &= (-T^3 + 4T^2 - 5T + 2I) + (10T - 4I) \\ &= p_T(T) + 10T - 4I = 10T - 4I \\ &= \begin{bmatrix} 10 & 0 & 20 \\ 0 & 10 & 10 \\ 0 & 0 & 20 \end{bmatrix} - \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 0 & 20 \\ 0 & 6 & 10 \\ 0 & 0 & 16 \end{bmatrix}. \end{aligned}$$

Therefore the answer is

$$-T^3 + 4T^2 + 5T - 2I = \begin{bmatrix} 6 & 0 & 20 \\ 0 & 6 & 10 \\ 0 & 0 & 16 \end{bmatrix}.$$

## Comment.

This problem is actually one of the final exam problems in the linear algebra class I taught.
Of course, you may calculate them directly but most students who computed directly made calculation mistakes (unfortunately).

The second common mistake is that students jumped to conclusion that the expression $-T^3 + 4T^2 + 5T - 2I$ is the zero matrix after finding the characteristic polynomial.
Those students knew that they could use the Cayley-Hamilton theorem but were a bit careless. The given expression and the characteristic polynomial are slightly different.

Anyway, I didn't subtract many points for the second mistake.

(And you get little points if you made a mistake for a direct calculation.)

②

Find the inverse matrix of the matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 9 & 2 & 0 \\ 5 & 0 & 3 \end{bmatrix}$ using the Cayley–Hamilton theorem.

## Solution.

To use the Cayley-Hamilton theorem, we first compute the characteristic polynomial $p(t)$ of the matrix $A$. We have

$$\begin{aligned} & p(t) = \det(A - tI) \\ & \begin{vmatrix} 1-t & 1 & 2 \\ 9 & 2-t & 0 \\ 5 & 0 & 3-t \end{vmatrix} \\ & = (-1)^{3+1} 5 \begin{vmatrix} 1 & 2 \\ 2-t & 0 \end{vmatrix} + (-1)^{3+2} \cdot 0 \begin{vmatrix} 1-t & 2 \\ 9 & 0 \end{vmatrix} + (-1)^{3+3}(3-t) \begin{vmatrix} 1-t & 1 \\ 9 & 2-t \end{vmatrix} \\ & \text{(by the 3rd row cofactor expansion)} \\ & = 5(2t-4) + 0 + (3-t)\left((1-t)(2-t) - 9\right) \\ & = -t^3 + 6t^2 + 8t - 41. \end{aligned}$$

Then the Cayley-Hamilton theorem yields that $p(A) = O$, the zero matrix. That is, we have

$$O = p(A) = -A^3 + 6A^2 + 8A - 41I.$$

Thus, we have

$$41I = -A^3 + 6A^2 + 8A = A(-A^2 + 6A + 8I),$$

or equivalently

$$I = A\left(\frac{1}{41}(-A^2 + 6A + 8I)\right).$$

It follows that the inverse matrix is given by

$$A^{-1} = \frac{1}{41}(-A^2 + 6A + 8I).$$

By a direct computation, we have

$$A^2 = \begin{bmatrix} 20 & 3 & 8 \\ 27 & 13 & 18 \\ 20 & 5 & 19 \end{bmatrix}$$

and

$$-A^2+6A+8I=-\begin{bmatrix}20&3&8\\27&13&18\\20&5&19\end{bmatrix}+6\begin{bmatrix}1&1&2\\9&2&0\\5&0&3\end{bmatrix}+8\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$
$$=\begin{bmatrix}-6&3&4\\27&7&-18\\10&-5&7\end{bmatrix}.$$

Therefore the inverse matrix is

$$A^{-1}=\frac{1}{41}\begin{bmatrix}-6&3&4\\27&7&-18\\10&-5&7\end{bmatrix}.$$

## More Exercise

Test whether you understand how to find the inverse matrix using the Cayley-Hamilton theorem by the next problem.

**Problem**. Find the inverse matrix of the $3 \times 3$ matrix

$$A=\begin{bmatrix}7&2&-2\\-6&-1&2\\6&2&-1\end{bmatrix}$$

using the Cayley-Hamilton theorem.

③ Find the inverse matrix of the $3 \times 3$ matrix $A=\begin{bmatrix}7&2&-2\\-6&-1&2\\6&2&-1\end{bmatrix}$ using the Cayley-Hamilton theorem.

## Solution.

To apply the Cayley-Hamilton theorem, we first determine the characteristic polynomial $p(t)$ of the matrix $A$.

Let $I$ be the $3 \times 3$ identity matrix.

We have

$$p(t)=\det(A-tI)$$

$$=\begin{vmatrix}7-t&2&-2\\-6&-1-t&2\\6&2&-1-t\end{vmatrix}$$

$$=(7-t)\begin{vmatrix}-1-t&2\\2&-1-t\end{vmatrix}-2\begin{vmatrix}-6&2\\6&-1-t\end{vmatrix}+(-2)\begin{vmatrix}-6&-1-t\\6&2\end{vmatrix}$$

(by the first row cofactor expansion)

$$=-t^3+5t^2-7t+3.$$

(You may also use the rule of Sarrus to compute the $3 \times 3$ determinant.)

Thus, we have obtained the characteristic polynomial

$$p(t) = -t^3 + 5t^2 - 7t + 3$$

of the matrix $A$.

The Cayley-Hamilton theorem yields that

$$O = p(A) = -A^3 + 5A^2 - 7A + 3I,$$

where $O$ is the $3 \times 3$ zero matrix.
(Here, don't forget to put the identity matrix $I$.)

Rearranging terms, we have

$$\begin{aligned} & A^3 - 5A^2 + 7A = 3I \\ & \Leftrightarrow A(A^2 - 5A + 7I) = 3I \\ & \Leftrightarrow A\left(\frac{1}{3}(A^2 - 5A + 7I)\right) = I. \end{aligned}$$

Similarly, we have

$$\left(\frac{1}{3}(A^2 - 5A + 7I)\right)A = I.$$

It follows from these two equalities that the matrix

$$\frac{1}{3}(A^2 - 5A + 7I)$$

is the inverse matrix of $A$.

Therefore, we have

$$\begin{aligned} A^{-1} &= \frac{1}{3}(A^2 - 5A + 7I) \\ &= \frac{1}{3}\left(\begin{bmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{bmatrix} - 5\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix} + 7\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\ &= \frac{1}{3}\begin{bmatrix} -3 & -2 & 2 \\ 6 & 5 & -2 \\ -6 & -2 & 5 \end{bmatrix}. \end{aligned}$$

In summary, the inverse matrix of $A$ is

$$A^{-1} = \frac{1}{3}\begin{bmatrix} -3 & -2 & 2 \\ 6 & 5 & -2 \\ -6 & -2 & 5 \end{bmatrix}.$$

4 Let

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}.$$

Find the eigenvalues and the eigenvectors of the matrix

$$B = A^4 - 3A^3 + 3A^2 - 2A + 8E.$$

(*Nagoya University*)

## Hint.

Apply the Cayley-Hamilton theorem.
That is if $p_A(t)$ is the characteristic polynomial of the matrix $A$, then the matrix $p_A(A)$ is the zero matrix.

## Solution.

Let us first find the characteristic polynomial $p_A(t)$ of the matrix $A$.
We have

$$\begin{aligned} p_A(t) = & \det(A - tI) = \begin{vmatrix} 1-t & -1 \\ 2 & 3-t \end{vmatrix} \\ & = (1-t)(3-t) - (-1)(2) = t^2 - 4t + 5. \end{aligned}$$

Solving $t^2 - 4t + 5 = 0$, we see that the matrix $A$ has the eigenvalues $2 \pm i$ but it is not a good idea to use this directly to find the eigenvalues of the matrix $B$.
Instead, note that by the Cayley-Hamilton theorem, we know that

$$p_t(A) = A^2 - 4A + 5I = O,$$

where $I$ is the $2 \times 2$ identity matrix and $O$ is the $2 \times 2$ zero matrix.

---

Since we have

$$B = A^4 - 3A^3 + 3A^2 - 2A + 8E = (A^2 - 4A + 5I)(A^2 + A + 2I) + A - 2I,$$

we have

$$B = A - 2I = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}.$$

Since the eigenvalues of $A$ is $2 \pm i$, the eigenvalues of $B = A - 2I$ are

$$(2 \pm i) - 2 = \pm i.$$

---

Next, we find eigenvectors.
Let us first find eigenvectors corresponding to the eigenvalue $i$.
We have

$$A - iI = \begin{bmatrix} -1-i & -1 \\ 2 & 1-i \end{bmatrix} \xrightarrow{(-1+i)R_1} \begin{bmatrix} 2 & 1-i \\ 2 & 1-i \end{bmatrix}$$

$$\xrightarrow{R_2-R_1} \begin{bmatrix} 2 & 1-i \\ 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & (1-i)/2 \\ 0 & 0 \end{bmatrix}.$$

Thus we have

$$x_1 = -\frac{1-i}{2}$$

and the eigenvectors associated with the eigenvalue $i$ are

$$\mathbf{x} = x_2 \begin{bmatrix} -\frac{1-i}{2} \\ 1 \end{bmatrix},$$

where $x_2$ is any nonzero complex number.
Or equivalently, scaling the vector by $-1+i$, the eigenvectors corresponding to the eigenvalue $i$ are

$$a \begin{bmatrix} 1 \\ -1-i \end{bmatrix},$$

where $a$ is any nonzero complex number.

---

Since $B$ is a real matrix and the eigenvalues $i$ and $-i$ are complex conjugate to each other, the eigenvectors of $-i$ are just the conjugates of eigenvectors of $i$. Thus the eigenvectors corresponding to the eigenvalue $-i$ are

$$b \begin{bmatrix} 1 \\ -1+i \end{bmatrix},$$

where $b$ is any nonzero complex number.

⑤ Let $A, B$ be complex $2 \times 2$ matrices satisfying the relation $A = AB - BA$. Prove that $A^2 = O$, where $O$ is the $2 \times 2$ zero matrix.

## Hint.

⑥ Find the trace of $A$ .

⑦ Use the Cayley-Hamilton theorem

## Proof.

We first calculate the trace of the matrix $A$ as follows. We have

$$\begin{aligned}\operatorname{tr}(A) &= \operatorname{tr}(AB-BA)\\ &= \operatorname{tr}(AB)-\operatorname{tr}(BA)\\ &= \operatorname{tr}(AB)-\operatorname{tr}(AB)=0.\end{aligned}$$

Thus $\operatorname{tr}(A)=0$ and it follows from the Cayley-Hamilton theorem (see below) for the $2\times 2$ matrix $A$ that

$$\begin{aligned}O &= A^2-\operatorname{tr}(A)A+\det(A)I\\ &= A^2+\det(A)I,\end{aligned}$$

where $I$ is the $2\times 2$ identity matrix.

Thus, we obtain

$$A^2=-\det(A)I. \tag{*}$$

---

Next, we compute $A^2$ in two ways.
We have

$$A^2=A(AB-BA)=A^2B-ABA$$

and

$$A^2=(AB-BA)A=ABA-BA^2.$$

Adding these two, we have

$$\begin{aligned}2A^2 &= A^2B-BA^2\\ &\overset{(*)}{=} (-\det(A)I)B-B(-\det(A)I)\\ &= -\det(A)B+\det(A)B=O.\end{aligned}$$

As a result, we obtain $A^2=O$. This completes the proof.

## The Cayley-Hamilton theorem for a $2\times 2$ matrix

Let us add the proof of the fact we used in the proof about the **Cayley-Hamilton theorem**.
Let $A=\begin{bmatrix} a & b\\ c & d\end{bmatrix}$ be a $2\times 2$ matrix.

Then its characteristic polynomial is

$$\begin{aligned}p(x) &= \det(A-xI)\\ &= \begin{vmatrix} a-x & b\\ c & d-x\end{vmatrix}\\ &= (a-x)(d-x)-bc\\ &= x^2-(a+d)x+ad-bc\\ &= x^2-\operatorname{tr}(A)x+\det(A),\end{aligned}$$

since $\mathrm{tr}(A) = a + d$ and $\det(A) = ad - bc$.

The Cayley-Hamilton theorem says that the matrix $A$ satisfies its characteristic equation $p(x) = 0$.
Namely we have

$$A^2 - \mathrm{tr}(A)A + \det(A)I = O.$$

This is the equality we used in the proof.

## Variation

As a variation of this problem, consider the following problem.

Let $A, B$ be $2 \times 2$ matrices satisfying $A = AB - BA$ .
Then prove that $\det(A) = 0$.

⑧ Suppose that $A$ is $2 \times 2$ matrix that has eigenvalues $-1$ and $3$. Then for each positive integer $n$ find $a_n$ and $b_n$ such that $A^{n+1} = a_n A + b_n I$, where $I$ is the $2 \times 2$ identity matrix.

## Solution.

Since $-1, 3$ are eigenvalues of the matrix $A$, the characteristic polynomial of $A$ is

$$(t+1)(t-3) = t^2 - 2t - 3.$$

By Cayley-Hamilton theorem, we have

$$A^2 - 2A - 3I = O,$$

where $O$ is the $2 \times 2$ zero matrix.

Thus we have

$$A^2 = 2A + 3I,$$

and hence $a_1 = 2, b_1 = 3.$

Multiplying (*) by $A$, we have

$$\begin{aligned} A^3 &= AA^2 = 2A^2 + 3A \\ &= 2(2A + 3I) + 3A \\ &= 7A + 6I. \end{aligned}$$

Thus, $a_2 = 7, b_2 = 6.$

In general, we have

$$\begin{aligned} A^{n+2} &= AA^{n+1} \\ &= A(a_n A + b_n I) \\ &= a_n A^2 + b_n A \\ &= a_n(2A + 3I) + b_n A \\ &= (2a_n + b_n)A + (3a_n)I. \end{aligned}$$

Thus we obtain

$$\begin{aligned} a_{n+1} &= 2a_n + b_n \\ b_{n+1} &= 3a_n. \end{aligned}$$

This yields the linear recurrence relation

$$a_{n+2} = 2a_{n+1} + 3a_n$$

with initial values $a_1 = 2, a_2 = 7$.

The general term $a_n$ of the sequence $(a_n)_{i=n}^{\infty}$ is obtained in the previous post Solve a linear recurrence relation using vector space technique:

$$a_n = \frac{1}{4}\big((-1)^n + 3^{n+1}\big).$$

(You may alternatively find this using an elementary technique.)

Then we also have

$$b_n = 3a_{n-1} = \frac{3}{4}\big((-1)^{n-1} + 3^n\big).$$

(This formula is true for $n = 1$ as well.)

In conclusion, we have obtained

$$A^n = \frac{1}{4}\big((-1)^n + 3^{n+1}\big)A + \frac{3}{4}\big((-1)^{n-1} + 3^n\big)I$$

for any positive integer $n$.

---

9 Let $A$ be an $n \times n$ complex matrix. Let $p(x) = \det(xI - A)$ be the characteristic polynomial of $A$ and write it as

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

where $a_i$ are real numbers. Let $C$ be the companion matrix of the polynomial $p(x)$ given by

$$C = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{bmatrix} = [\mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n, -\mathbf{a}],$$

where $\mathbf{e}_i$ is the unit vector in $\mathbb{C}^n$ whose $i$-th entry is $1$ and zero elsewhere, and the vector $\mathbf{a}$ is defined by $\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix}$. Then prove that the following two statements are equivalent.

(1) There exists a vector $\mathbf{v} \in \mathbb{C}^n$ such that

$$\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^{n-1}\mathbf{v}$$

form a basis of $\mathbb{C}^n$.
(2) There exists an invertible matrix $S$ such that $S^{-1}AS = C$.
(Namely, $A$ is similar to the companion matrix of its characteristic polynomial.)

## Proof.

Before proving the equivalence of the statements, let us first note a simple but useful fact.
Let $B$ be $n \times n$ matrix and let $\mathbf{e}_i$ be the unit vector of $\mathrm{C}^n$ as defined in the problem. Then the matrix product $B\mathbf{e}_i$ is the $i$-th column vector of the matrix $B$.

### $(1 \implies 2)$

Suppose that the statement 1 is true. That is, there exists a vector $\mathbf{v} \in \mathrm{C}^n$ such that

$$\mathbf{v}, A\mathbf{v}, \dots, A^{n-1}\mathbf{v}$$

form a basis of $\mathrm{C}^n$. In particular, these vectors are linearly independent.

Form a matrix

$$S := [\mathbf{v}, A\mathbf{v}, \dots, A^{n-1}\mathbf{v}],$$

then the matrix $S$ is invertible since the columns vectors are linearly independent.

We prove that $AS = SC$.
We have

$$\begin{aligned} AS &= A[\mathbf{v}, A\mathbf{v}, \dots, A^{n-1}\mathbf{v}], \\ &= [A\mathbf{v}, A^2\mathbf{v}, \dots, A^n\mathbf{v}]. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} SC &= S[\mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n, -\mathbf{a}] \\ &= [S\mathbf{e}_2, S\mathbf{e}_3, \dots, S\mathbf{e}_n, -S\mathbf{a}]. \end{aligned}$$

As noted at the beginning, the vector $S\mathbf{e}_2$ is the second column vector of $S$, which is $A\mathbf{v}$. Similarly, $S\mathbf{e}_i$ is the $i$-th column vector of $S$, and thus we have

$$S\mathbf{e}_i = A^{i-1}\mathbf{v}.$$

Hence we have

$$SC = [A\mathbf{v}, A^2\mathbf{v}, \dots, A^{n-1}\mathbf{v}, -S\mathbf{a}].$$

Thus, it remains to prove that $-S\mathbf{a} = A^n\mathbf{v}$.
Write

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = a_0\mathbf{e}_1 + a_1\mathbf{e}_2 + \cdots + a_{n-1}\mathbf{e}_n.$$

Then we have

$$\begin{aligned} S\mathbf{a} &= S(a_0\mathbf{e}_1 + a_1\mathbf{e}_2 + \cdots + a_{n-1}\mathbf{e}_n) \\ &= a_0S\mathbf{e}_1 + a_1S\mathbf{e}_2 + \cdots + a_{n-1}S\mathbf{e}_n \\ &\overset{(*)}{=} a_0\mathbf{v} + a_1A\mathbf{v} + \cdots + a_{n-1}A^{n-1}\mathbf{v} \\ &= (a_0I + a_1A + \cdots + a_{n-1}A^{n-1})\mathbf{v} \\ &= -A^n\mathbf{v}, \end{aligned}$$

where the last step follows from the Cayley-Hamilton theorem

$$O = p(A) = A^n + a_{n-1}A^{n-1} + \cdots a_1A + a_0I.$$

This proves that $-S\mathbf{a} = A^n\mathbf{v}$ and hence $AS = SC$.
Since $S$ is invertible, we obtain

$$S^{-1}AS = C.$$

### (2 $\implies$ 1)

Next suppose that the statement 2 is true. That is, there exists an invertible matrix $S$ such that $S^{-1}AS = C$. Then we have $AS = CS$.
Define $\mathbf{v} := S\mathbf{e}_1$ to be the first column vector of $S$.

We prove that with this choice of $\mathbf{v}$, the vectors

$$\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$$

form a basis of $\mathrm{C}^n$.

We have

$$\begin{aligned} AS &= A[S\mathbf{e}_1, S\mathbf{e}_2, \ldots, S\mathbf{e}_n] \\ &= [AS\mathbf{e}_1, AS\mathbf{e}_2, \ldots, AS\mathbf{e}_n] \end{aligned}$$

and

$$\begin{aligned} SC &= S[\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n, -\mathbf{a}] \\ &= [S\mathbf{e}_2, S\mathbf{e}_3, \ldots, S\mathbf{e}_n, -S\mathbf{a}]. \end{aligned}$$

Since $AS = SC$, comparing column vectors, we obtain

$$\begin{aligned} AS\mathbf{e}_1 &= S\mathbf{e}_2 \\ AS\mathbf{e}_2 &= S\mathbf{e}_3 \\ &\vdots \\ AS\mathbf{e}_{n-1} &= S\mathbf{e}_n \\ AS\mathbf{e}_n &= -S\mathbf{a}. \end{aligned}$$

It follows that we have inductively

$$\begin{aligned} A\mathbf{v} &= S\mathbf{e}_2 \\ A^2\mathbf{v} &= A(A\mathbf{v}) = AS\mathbf{e}_2 = S\mathbf{e}_3 \\ &\vdots \\ A^{n-1}\mathbf{v} &= A(A^{n-2}\mathbf{v}) = AS\mathbf{e}_{n-1} = S\mathbf{e}_n. \end{aligned}$$

Namely we obtain

$$A^i\mathbf{v} = S\mathbf{e}_{i+1}$$

for $i = 0, 1, \dots, n-1$.

Now, suppose that we have the linear combinaiton

$$c_1\mathbf{v} + c_2A\mathbf{v} + \dots c_nA^{n-1}\mathbf{v} = \mathbf{0},$$

for some coefficient scalars $c_1, \dots, c_n \in \mathbb{C}$.

Then we have using the above relations

$$c_1S\mathbf{e}_1 + c_2S\mathbf{e}_2 + \cdots + c_nS\mathbf{e}_n = \mathbf{0}.$$

Since $S$ is invertible, the column vectors $S\mathbf{e}_i$ are linearly independent.

Thus, we must have $c_1 = c_2 = \cdots = c_n = 0$.

Therefore the vectors

$$\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^{n-1}\mathbf{v}$$

are linearly independent, and they form a basis of $\mathbb{C}^n$ since we have $n$ linearly independent vectors in the $n$-dimensional vector space $\mathbb{C}^n$.

---

**10** Let $n > 1$ be a positive integer. Let $V = M_{n\times n}(\mathbb{C})$ be the vector space over the complex numbers $\mathbb{C}$ consisting of all complex $n \times n$ matrices. The dimension of $V$ is $n^2$. Let $A \in V$ and consider the set

$$S_A = \{I = A^0, A, A^2, \dots, A^{n^2-1}\}$$

of $n^2$ elements. Prove that the set $S_A$ cannot be a basis of the vector space $V$ for any $A \in V$.

## Proof.

We prove that the set $S_A$ is linearly dependent, hence it cannot be a basis of $V$.

Since $A$ is an $n \times n$ matrix, its characteristic polynomial $p(t) = \det(tI - A)$ is a degree $n$ polynomial.

(Your preferred definition of the characteristic polynomial might be $\det(A - tI)$. It is straight forward to modify the following proof with this definition.)

Let us write it as

$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1x + a_0.$$

Then the Cayley-Hamilton theorem states that

$$p(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = O$$

is the zero matrix.

Since the coefficient of $A^n$ is $1$, this gives a non-trivial linear combination of $I, A, \dots, A^n$. Therefore the set

$$T := \{I, A, \dots, A^n\}$$

is linearly dependent.

As $T$ is a subset of $S_A$, the set $S_A$ is also linearly dependent.
Therefore, $S_A$ is not a basis of $V$. This completes the proof.

---

⑪ Let $A$ be a $3 \times 3$ real orthogonal matrix with $\det(A) = 1$.
**(a)** If $\frac{-1+\sqrt{3}i}{2}$ is one of the eigenvalues of $A$, then find the all the eigenvalues of $A$.
**(b)** Let $A^{100} = aA^2 + bA + cI$, where $I$ is the $3 \times 3$ identity matrix.
Using the Cayley-Hamilton theorem, determine $a, b, c$.
(*Kyushu University*)

## Solution.

### (a) Find the all the eigenvalues of $A$.

Since $A$ is a real matrix and $\frac{-1+\sqrt{3}i}{2}$ is a complex eigenvalue, its conjugate $\frac{-1-\sqrt{3}i}{2}$ is also an eigenvalue of $A$.
As $A$ is a $3 \times 3$ matrix, it has one more eigenvalue $\lambda$.

Note that the product of all eigenvalues of $A$ is the determinant of $A$.
Thus, we have

$$\frac{-1+\sqrt{3}i}{2} \cdot \frac{-1-\sqrt{3}i}{2} \cdot \lambda = \det(A) = 1.$$

Solving this, we obtain $\lambda = 1$.
Therefore, the eigenvalues of $A$ are

$$\frac{-1+\sqrt{3}i}{2}, \frac{-1-\sqrt{3}i}{2}, 1.$$

### (a) Using the Cayley-Hamilton theorem, determine $a, b, c$.

To use the Cayley-Hamilton theorem, we first need to determine the characteristic polynomial $p(t) = \det(A - tI)$ of $A$.
Since we found all the eigenvalues of $A$ in part (a) and the roots of characteristic polynomials are the eigenvalues, we know that

$$\begin{aligned} p(t) &= -\left(t - \frac{-1+\sqrt{3}i}{2}\right)\left(t - \frac{-1-\sqrt{3}i}{2}\right)(t-1) \qquad (*) \\ &= -(t^2+t+1)(t-1) \\ &= -t^3+1. \end{aligned}$$

(Remark that if your definition of the characteristic polynomial is $\det(tI-A)$, then the first negative sign in (*) should be omitted.)

Then the Cayley-Hamilton theorem yields that

$$P(A) = -A^3 + I = O,$$

where $O$ is the $3 \times 3$ zero matrix.

Hence we have $A^3 = I$.
We compute

$$A^{100} = (A^3)^{33}A = I^{33}A = IA = A.$$

Thus, we conclude that $a = 0, b = 1, c = 0$.

## Comment.

Observe that we did not use the assumption that $A$ is orthogonal.

---

12 Let $A$ and $B$ be $2 \times 2$ matrices such that $(AB)^2 = O$, where $O$ is the $2 \times 2$ zero matrix. Determine whether $(BA)^2$ must be $O$ as well. If so, prove it. If not, give a counter example.

## Proof.

It is true that the matrix $(BA)^2$ must be the zero matrix as we shall prove now.

For notational convenience, let $C := AB$.

As $C$ is a $2 \times 2$ matrix, it satisfies the relation

$$C^2 - \operatorname{tr}(C)C + \det(C)I = O \qquad (*)$$

by the Cayley-Hamilton theorem.
Here $I$ is the $2 \times 2$ identity matrix.

---

We compute the determinant of $C$ as follows.
We have

$$\det(C)^2 = \det(C^2) = \det((AB)^2) = \det(O) = 0.$$

Hence $\det(C) = 0$.

Since $C^2 = O$ and $\det(C) = 0$, the Cayley-Hamilton relation (*) becomes

$$\operatorname{tr}(C)C = O.$$

It follows that we have either $C = O$ or $\operatorname{tr}(C) = 0$.
In either case, we have $\operatorname{tr}(C) = 0$.

---

Note that

$$\begin{gathered}\det(BA) = \det(AB) = \det(C) = 0 \\ \operatorname{tr}(BA) = \operatorname{tr}(AB) = \operatorname{tr}(C) = 0.\end{gathered}$$

Applying the Cayley-Hamilton theorem to the matrix $BA$, we obtain

$$O = (BA)^2 - \operatorname{tr}(BA) \cdot BA + \det(BA)I = (BA)^2.$$

Thus, we obtain $(BA)^2 = O$ as claimed.

# Dot Products and Length of Vectors

**Definition**

Let $A$ be a square matrix.

① Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are **orthogonal** (**perpendicular**) if $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

② Let $\mathbf{v}$ be $n$-dimensional complex vector. Then the **length** of $\mathbf{v}$ is defined to be

$$\|\mathbf{v}\| = \sqrt{\bar{\mathbf{v}}^{\mathrm{T}}\mathbf{v}}.$$

③ The **distance** between two vectors $\mathbf{v}_1, \mathbf{v}_2$ is the length $\|\mathbf{v}_1 - \mathbf{v}_2\|$.

④ $A$ is **orthogonal** if $A^{\mathrm{T}}A = I = AA^{\mathrm{T}}$.

⑤ $A$ is **normal** if $A^*A = A = AA^*$. Here $A^* = \bar{A}^{\mathrm{T}}$.

**Summary**

Let $A$ be a square matrix. Let $\mathbf{v} \in K^n$, where $K = \mathbb{R}$ or $\mathbb{C}$.

① $\|\mathbf{v}\| \geq 0$.

② $A$ is orthogonal if and only if its column vectors form an orthonormal set.

## Problems

① Let $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are vectors in $\mathbb{R}^n$. Suppose that vectors $\mathbf{u}_1$, $\mathbf{u}_2$ are orthogonal and the norm of $\mathbf{u}_2$ is $4$ and $\mathbf{u}_2^{\mathrm{T}}\mathbf{u}_3 = 7$. Find the value of the real number $a$ in $\mathbf{u_1} = \mathbf{u_2} + a\mathbf{u}_3$.
(*The Ohio State University*)

**Hint.**

Recall the following definitions.

② The inner product (dot product) of two vectors $\mathbf{v}_1$, $\mathbf{v}_2$ is defined to be

$$\mathbf{v}_1 \cdot \mathbf{v}_2 := \mathbf{v}_1^{\mathrm{T}}\mathbf{v}_2.$$

③ Two vectors $\mathbf{v}_1$, $\mathbf{v}_2$ are orthogonal if the inner product

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

④ The norm (length, magnitude) of a vector **v** is defined to be

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

## Solution.

We first express the given conditions in term of inner products (dot products).

Since $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal, the inner product

$$\mathbf{u}_2 \cdot \mathbf{u}_1 = \mathbf{u}_1 \cdot \mathbf{u}_2 = 0.$$

Also, since the norm of $\mathbf{u}_2$ is 4, we obtain

$$\mathbf{u}_2 \cdot \mathbf{u}_2 = ||\mathbf{u}_2||^2 = 16.$$

The last condition can be written as

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = \mathbf{u}_2^{\mathrm{T}}\mathbf{u}_3 = 7.$$

Now we compute the inner product $\mathbf{u}_2$ and $\mathbf{u}_1$.

$$\begin{aligned} 0 &\overset{(a)}{=} \mathbf{u}_2 \cdot \mathbf{u}_1 = \mathbf{u}_2 \cdot (\mathbf{u}_2 + a\mathbf{u}_3) \\ &= \mathbf{u}_2 \cdot \mathbf{u}_2 + a\mathbf{u}_2 \cdot \mathbf{u}_3 \\ &\overset{(b),(c)}{=} 16 + 7a. \end{aligned}$$

Therefore, solving this we obtain

$$a = -\frac{16}{7}.$$

---

⑤ Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{R}^n$ such that their length are $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$ and the inner product

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b} = -\frac{1}{2}.$$

Then determine the length $\|\mathbf{a} - \mathbf{b}\|$.

## Solution.

Recall that the length of a vector **x** is defined to be

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}},$$

where $\mathbf{x}^{\mathrm{T}}$ is the transpose of **x**.

Also, recall that the inner product of two vectors **x**, **y** are commutative.

Namely we have

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{x} = \mathbf{y} \cdot \mathbf{x}.$$

Applying the second fact with given vectors $\mathbf{a}$, $\mathbf{b}$, we obtain

$$\mathbf{a}^{\mathrm{T}}\mathbf{b} = \mathbf{b}^{\mathrm{T}}\mathbf{a} = -\frac{1}{2}.$$

Now we compute $\|\mathbf{a} - \mathbf{b}\|^2$ as follows.

We have

$$\begin{aligned}
\|\mathbf{a} - \mathbf{b}\|^2 &= (\mathbf{a} - \mathbf{b})^{\mathrm{T}}(\mathbf{a} - \mathbf{b}) \qquad \text{(by definition of the length)} \\
&= (\mathbf{a}^{\mathrm{T}} - \mathbf{b}^{\mathrm{T}})(\mathbf{a} - \mathbf{b}) \\
&= \mathbf{a}^{\mathrm{T}}\mathbf{a} - \mathbf{a}^{\mathrm{T}}\mathbf{b} - \mathbf{b}^{\mathrm{T}}\mathbf{a} + \mathbf{b}^{\mathrm{T}}\mathbf{b} \\
&= \|\mathbf{a}\|^2 - \mathbf{a}^{\mathrm{T}}\mathbf{b} - \mathbf{b}^{\mathrm{T}}\mathbf{a} + \|\mathbf{b}\|^2 \\
&= 1 - \left(-\frac{1}{2}\right) - \left(-\frac{1}{2}\right) + 1 \\
&= 3.
\end{aligned}$$

Since the length is nonnegative, we take the square root of the above equality and obtain

$$\|\mathbf{a} - \mathbf{b}\| = \sqrt{3}.$$

⑥ Let $A$ and $B$ be $n \times n$ skew-symmetric matrices. Namely $A^{\mathrm{T}} = -A$ and $B^{\mathrm{T}} = -B$.
**(a)** Prove that $A + B$ is skew-symmetric.
**(b)** Prove that $cA$ is skew-symmetric for any scalar $c$.
**(c)** Let $P$ be an $m \times n$ matrix. Prove that $P^{\mathrm{T}}AP$ is skew-symmetric.
**(d)** Suppose that $A$ is real skew-symmetric. Prove that $iA$ is an Hermitian matrix.
**(e)** Prove that if $AB = -BA$, then $AB$ is a skew-symmetric matrix.
**(f)** Let $\mathbf{v}$ be an $n$-dimensional column vecotor. Prove that $\mathbf{v}^{\mathrm{T}}A\mathbf{v} = 0$.
**(g)** Suppose that $A$ is a real skew-symmetric matrix and $A^2\mathbf{v} = \mathbf{0}$ for some vector $\mathbf{v} \in \mathbb{R}^n$. Then prove that $A\mathbf{v} = \mathbf{0}$.

## Proof.

### (a) Prove that $A + B$ is skew-symmetric.

We have

$$(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}} = (-A) + (-B) = -(A + B).$$

Hence $A + B$ is skew-symmetric.

### (b) Prove that $cA$ is skew-symmetric for any scalar $c$.

We compute

$$(cA)^{\mathrm{T}} = cA^{\mathrm{T}} = c(-A) = -cA.$$

Thus, $cA$ is skew-symmetric.

### (c) Let $P$ be an $m \times n$ matrix. Prove that $P^{\mathrm{T}}AP$ is skew-symmetric.

Using the properties of transpose, we have

$$\begin{aligned}(P^{\mathrm{T}}AP)^{\mathrm{T}} &= P^{\mathrm{T}}A^{\mathrm{T}}(P^{\mathrm{T}})^{\mathrm{T}} = P^{\mathrm{T}}A^{\mathrm{T}}P \\ &= P^{\mathrm{T}}(-A)P = -(P^{\mathrm{T}}AP).\end{aligned}$$

This implies that $P^{\mathrm{T}}AP$ is skew-symmetric.

(d) Suppose that $A$ is real skew-symmetric. Prove that $iA$ is an Hermitian matrix.

Note that since $A$ is real, we have $\bar{A} = A$.

Then we have

$$\overline{(iA)}^{\mathrm{T}} = (\bar{i}\bar{A})^{\mathrm{T}} = (-iA)^{\mathrm{T}} = (-i)A^{\mathrm{T}} = (-i)(-A) = iA.$$

It follows that $iA$ is Hermitian.

(e) Prove that if $AB = -BA$, then $AB$ is a skew-symmetric matrix.

We calculate

$$\begin{aligned}(AB)^{\mathrm{T}} &= B^{\mathrm{T}}A^{\mathrm{T}} = (-B)(-A) \\ &= BA = -AB,\end{aligned}$$

where the last step follows from the assumption $AB = -BA$.

This proves that $AB$ is skew-symmetric.

(f) Let $\mathbf{v}$ be an $n$-dimensional column vecotor. Prove that $\mathbf{v}^{\mathrm{T}}A\mathbf{v} = 0$.

Observe that $\mathbf{v}^{\mathrm{T}}A\mathbf{v}$ is a $1 \times 1$ matrix, or just a number.

So we have

$$\begin{aligned}\mathbf{v}^{\mathrm{T}}A\mathbf{v} &= (\mathbf{v}^{\mathrm{T}}A\mathbf{v})^{\mathrm{T}} = \mathbf{v}^{\mathrm{T}}A^{\mathrm{T}}(\mathbf{v}^{\mathrm{T}})^{\mathrm{T}} \\ &= \mathbf{v}^{\mathrm{T}}A^{\mathrm{T}}\mathbf{v} = \mathbf{v}^{\mathrm{T}}(-A)\mathbf{v} = -(\mathbf{v}^{\mathrm{T}}A\mathbf{v}).\end{aligned}$$

This yields that $2\mathbf{v}^{\mathrm{T}}A\mathbf{v} = 0$, and hence $\mathbf{v}^{\mathrm{T}}A\mathbf{v} = 0$.

(g) Suppose that $A$ is a real skew-symmetric matrix and $A^2\mathbf{v} = \mathbf{0}$ for some vector $\mathbf{v} \in \mathrm{R}^n$. Then prove that $A\mathbf{v} = \mathbf{0}$.

Let us compute the length of the vector $A\mathbf{v}$.

We have

$$\begin{aligned}\|A\mathbf{v}\| &= (A\mathbf{v})^{\mathrm{T}}(A\mathbf{v}) = \mathbf{v}^{\mathrm{T}}A^{\mathrm{T}}A\mathbf{v} \\ &= \mathbf{v}^{\mathrm{T}}(-A)A\mathbf{v} = -\mathbf{v}^{\mathrm{T}}A^2\mathbf{v} \\ &= -\mathbf{v}\mathbf{0} && \text{by assumption} \\ &= 0.\end{aligned}$$

Since the length $\|A\mathbf{v}\| = 0$, we conclude that $A\mathbf{v} = \mathbf{0}$.

7 Let $\mathbf{v}$ be a vector in an inner product space $V$ over $\mathbb{R}$. Suppose that $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal basis of $V$. Let $\theta_i$ be the angle between $\mathbf{v}$ and $\mathbf{u}_i$ for $i = 1, \ldots, n$. Prove that

$$\cos^2\theta_1 + \cdots + \cos^2\theta_n = 1.$$

## Definition (Angle between Vectors).

Let $\langle \mathbf{a}, \mathbf{b} \rangle$ denote the inner product of vectors $\mathbf{a}$ and $\mathbf{b}$ in $V$.

Recall that the angle $\theta$ between $\mathbf{a}$ and $\mathbf{b}$ is defined as the unique number $\theta$ between 0 and $\pi$ satisfying

$$\cos\theta = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\|\|\mathbf{b}\|}.$$

## Proof.

Express the vector $\mathbf{v}$ as a linear combination of the basis vectors as

$$\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n$$

for some real numbers $a_1, \ldots, a_n$.

The length of the vector $\mathbf{v}$ is given by

$$\|\mathbf{v}\| = \sqrt{a_1^2 + \cdots + a_n^2}.$$

For each $i$, we have using the properties of the inner product

$$\begin{aligned}\langle \mathbf{v}, \mathbf{u}_i \rangle &= \langle a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n, \mathbf{u}_i \rangle \\ &= a_1\langle \mathbf{u}_1, \mathbf{u}_i \rangle + \cdots + a_n\langle \mathbf{u}_n, \mathbf{u}_i \rangle \\ &= a_i\end{aligned}$$

since $\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 1$ and $\langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0$ if $j \neq i$ as $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is orthonormal.

By definition of the angle, we have

$$\cos\theta_i = \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\|\mathbf{v}\|\|\mathbf{u}_i\|} = \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\|\mathbf{v}\|} \qquad \text{since } \|\mathbf{u}_i\| = 1.$$

It follows that

$$\begin{aligned}\cos^2\theta_1 + \cdots + \cos^2\theta_n &= \frac{\langle \mathbf{v}, \mathbf{u}_1\rangle^2}{\|\mathbf{v}\|^2} + \cdots + \frac{\langle \mathbf{v}, \mathbf{u}_n\rangle^2}{\|\mathbf{v}\|^2} \\ &= \frac{1}{\|\mathbf{v}\|^2}(a_1^2 + \cdots a_n^2) && \text{by } (**) \\ &= \frac{1}{\|\mathbf{v}\|^2} \cdot \|\mathbf{v}\|^2 && \text{by } (*) \\ &= 1.\end{aligned}$$

Thus we obtain

$$\cos^2\theta_1 + \cdots + \cos^2\theta_n = 1$$

as required.

---

8 Suppose that $A$ is a real $n \times n$ matrix.
**(a)** Is it true that $A$ must commute with its transpose?
**(b)** Suppose that the columns of $A$ (considered as vectors) form an orthonormal set. Is it true that the rows of $A$ must also form an orthonormal set?
(*University of California, Berkeley*)

## Solution.

### (a) Is it true that $A$ must commute with its transpose?

The answer is no.

We give a counterexample. Let

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

Then the transpose of $A$ is

$$A^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}.$$

We compute

$$AA^{\mathrm{T}} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix},$$

and

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}.$$

Therefore, we see that

$$AA^{\mathrm{T}} \neq A^{\mathrm{T}}A,$$

that is, $A$ does not commute with its transpose $A^{\mathrm{T}}$.

### (b) Is it true that the rows of $A$ must also form an orthonormal set?

The answer is yes.

Note that in general the column vectors of a matrix $M$ form an orthonormal set if and only if $M^{\mathrm{T}}M = I$, where $I$ is the identity matrix. (Such a matrix is called orthogonal matrix.)

Thus, by assumption we have $A^{\mathrm{T}}A = I$. Let $B = A^{\mathrm{T}}$.
Then the column vectors of $B$ is the row vectors of $A$. Hence it suffices to show that $B^{\mathrm{T}}B = I$.

Since $A^{\mathrm{T}}A = I$, we know that $A$ is invertible and the inverse $A^{-1} = A^{\mathrm{T}}$.
In particular, we have $A^{\mathrm{T}}A = AA^{\mathrm{T}} = I$.

We have

$$B^{\mathrm{T}}B = (A^{\mathrm{T}})^{\mathrm{T}}A^{\mathrm{T}} = (AA^{\mathrm{T}})^{\mathrm{T}} = I^{\mathrm{T}} = I.$$

Thus, we obtain $B^{\mathrm{T}}B = I$ and by the general fact stated above, the column vectors of $B$ form an orthonormal set.
Hence the row column vectors of $A$ form an orthonormal set.

9 Recall that a complex matrix is called **Hermitian** if $A^* = A$, where $A^* = \bar{A}^{\mathrm{T}}$. Prove that every Hermitian matrix $A$ can be written as the sum $A = B + iC$, where $B$ is a real symmetric matrix and $C$ is a real skew-symmetric matrix.

Proof.

Since $A$ is Hermitian, we have

$$\bar{A}^{\mathrm{T}} = A.$$

Taking the conjugate of this identity, we also have

$$A^{\mathrm{T}} = \bar{A}. \tag{*}$$

Let

$$B = \frac{1}{2}(A + \bar{A})$$

and

$$C = \frac{1}{2i}(A - \bar{A}).$$

We claim that $B$ is a real symmetric matrix and $C$ is a real skew-symmetric matrix.

We have

$$\bar{B} = \frac{1}{2}\overline{(A + \bar{A})} = \frac{1}{2}(\bar{A} + \bar{\bar{A}}) = \frac{1}{2}(\bar{A} + A) = B.$$

Thus, the matrix $B$ is real. To prove $B$ is symmetric, we compute

$$\begin{aligned}
B^{\mathrm{T}} &= \frac{1}{2}(A+\bar{A})^{\mathrm{T}} \\
&= \frac{1}{2}(A^{\mathrm{T}}+\bar{A}^{\mathrm{T}}) \\
&= \frac{1}{2}(\bar{A}+A) && \text{by (*) and } A \text{ is Hermitian} \\
&= B.
\end{aligned}$$

This proves that $B$ is symmetric.

The matrix $C$ is real because we have

$$\bar{C} = \frac{1}{-2i}(\bar{A}-\bar{\bar{A}}) = \frac{1}{2i}(A-\bar{A}) = C.$$

We also have

$$\begin{aligned}
C^{\mathrm{T}} &= \frac{1}{2i}(A^{\mathrm{T}}-\bar{A}^{\mathrm{T}}) \\
&= \frac{1}{2i}(\bar{A}-A) && \text{by (*) and } A \text{ is Hermitian} \\
&= -\frac{1}{2i}(A-\bar{A}) = -C.
\end{aligned}$$

Hence $C$ is a skew-symmetric matrix.

Finally, we compute

$$\begin{aligned}
B+iC &= \frac{1}{2}(A+\bar{A}) + i\cdot\frac{1}{2i}(A-\bar{A}) \\
&= \frac{1}{2}(A+\bar{A}) + \frac{1}{2}(A-\bar{A}) \\
&= A.
\end{aligned}$$

Therefore, we have obtained the sum as described in the problem.

---

10 **(a)** Prove that each complex $n \times n$ matrix $A$ can be written as $A = B + iC$, where $B$ and $C$ are Hermitian matrices.

**(b)** Write the complex matrix $A = \begin{bmatrix} i & 6 \\ 2-i & 1+i \end{bmatrix}$ as a sum $A = B + iC$, where $B$ and $C$ are Hermitian matrices.

### Definition (Hermitian matrix).

Recall that a complex matrix $M$ is said to be **Hermitian** if $M^* = M$.
Here $A^*$ is the conjugate transpose matrix $M^* = \bar{M}^*$.

### Proof.

Let

$$B=\frac{A+A^*}{2} \text{ and } C=\frac{A-A^*}{2i}.$$

We claim that $B$ and $C$ are Hermitian matrices.
Using the fact that $(A^*)^*=A$, we compute

$$\begin{aligned} B^* &= \left(\frac{A+A^*}{2}\right)^* \\ &= \frac{A^*+(A^*)^*}{2} \\ &= \frac{A^*+A}{2} = B. \end{aligned}$$

It yields that the matrix $B$ is Hermitian.

---

We also have

$$\begin{aligned} C^* &= \left(\frac{A-A^*}{2i}\right)^* \\ &= \frac{A^*-(A^*)^*}{-2i} \\ &= \frac{A^*-A}{-2i} \\ &= \frac{A-A^*}{2i} = C. \end{aligned}$$

Thus, the matrix $C$ is also Hermitian.

---

Finally, note that we have

$$\begin{aligned} B+iC &= \frac{A+A^*}{2}+i\frac{A-A^*}{2i} \\ &= \frac{A+A^*}{2}+\frac{A-A^*}{2} \\ &= A. \end{aligned}$$

Therefore, each complex matrix $A$ can be written as $A=B+iC$, where $B$ and $C$ are Hermitian matrices.

\item By the proof of part (a), it suffices to compute

$$B=\frac{A+A^*}{2} \text{ and } C=\frac{A-A^*}{2i}.$$

We have

$$A^*=\begin{bmatrix} -i & 2+i \\ 6 & 1-i \end{bmatrix}.$$

A direct computation yields that

$$B = \begin{bmatrix} 0 & 4 + \frac{i}{2} \\ 4 - \frac{i}{2} & 1 \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 & -\frac{1}{2} - 2i \\ -\frac{1}{2} + 2i & 1 \end{bmatrix}.$$

By the result of part (a), these matrices are Hermitian and satisfy $A = B + iC$, as required.

---

⑪ Let $\mathbf{a}, \mathbf{b}$ be vectors in $\mathbb{R}^n$. Prove the **Cauchy-Schwarz inequality**: $|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|$.

### Proof 1

Let $x$ be a variable and consider the length of the vector $\mathbf{a} - x\mathbf{b}$ as follows.

We have

$$\begin{aligned} 0 \leq \|\mathbf{a} - x\mathbf{b}\|^2 &= (\mathbf{a} - x\mathbf{b}) \cdot (\mathbf{a} - x\mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot x\mathbf{b} - x\mathbf{a} \cdot \mathbf{b} + x^2\mathbf{b} \cdot \mathbf{b} \\ &= \|\mathbf{b}\|^2 x^2 - 2\mathbf{a} \cdot \mathbf{b}x + \|\mathbf{a}\|^2. \end{aligned} \tag{*}$$

Note that the last expression is an equation of a parabola (quadratic equation).

Since the parabola is always non-negative, its discriminant $D$ must be non-positive.

Hence we have

$$D = (2\mathbf{a} \cdot \mathbf{b})^2 - 4\|\mathbf{a}\|^2 \|\mathbf{b}\|^2 \leq 0.$$

It follows that we have

$$(\mathbf{a} \cdot \mathbf{b})^2 \leq \|\mathbf{a}\|^2 \|\mathbf{b}\|^2.$$

Taking the square root, we obtain the Cauchy-Schwarz inequality

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|.$$

### Proof 2

The second proof starts with the same argument as the first proof.

As in Proof 1 (*), we obtain

$$0 \leq \|\mathbf{b}\|^2 x^2 - 2\mathbf{a} \cdot \mathbf{b}x + \|\mathbf{a}\|^2.$$

Now we take

$$x = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}.$$

Then we have

$$
\begin{aligned}
0 &\leq \|\mathbf{b}\|^2 \left( \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{b}\|^2} \right)^2 - 2\mathbf{a}\cdot\mathbf{b} \left( \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{b}\|^2} \right) + \|\mathbf{a}\|^2 \\
&= \frac{(\mathbf{a}\cdot\mathbf{b})^2}{\|\mathbf{b}\|^2} - 2\frac{(\mathbf{a}\cdot\mathbf{b})^2}{\|\mathbf{b}\|^2} + \|\mathbf{a}\|^2 \\
&= -\frac{(\mathbf{a}\cdot\mathbf{b})^2}{\|\mathbf{b}\|^2} + \|\mathbf{a}\|^2.
\end{aligned}
$$

It follows that we have

$$(\mathbf{a}\cdot\mathbf{b})^2 \leq \|\mathbf{a}\|^2 \|\mathbf{b}\|^2.$$

The Cauchy-Schwarz inequality is obtained by taking the square root as in Proof 1.

12 Let $A_1, A_2, \dots, A_m$ be $n \times n$ Hermitian matrices. Show that if

$$A_1^2 + A_2^2 + \cdots + A_m^2 = \mathcal{O},$$

where $\mathcal{O}$ is the $n \times n$ zero matrix, then we have $A_i = \mathcal{O}$ for each $i = 1, 2, \dots, m$.

### Hint.

Recall that a complex matrix $A$ is Hermitian if the conjugate transpose of $A$ is $A$ itself.

Namely, $A$ is Hermitian if

$$\bar{A}^{\mathrm{T}} = A.$$

We also use the length of a vector in the proof below.

Let $\mathbf{v}$ be $n$-dimensional complex vector. Then the length of $\mathbf{v}$ is defined to be

$$\|\mathbf{v}\| = \sqrt{\bar{\mathbf{v}}^{\mathrm{T}}\mathbf{v}}.$$

The length of a complex vector $\mathbf{v}$ is a non-negative real number.

The length is also called norm or magnitude.

### Proof.

Let $\mathbf{x}$ be an $n$-dimensional vector, that is, $\mathbf{x} \in \mathrm{R}^n$.

Then for each $i$, we have

$$\bar{\mathbf{x}}^{\mathrm{T}} A_i^2 \mathbf{x} = \bar{\mathbf{x}}^{\mathrm{T}} \bar{A}_i^{\mathrm{T}} A_i \mathbf{x} = (\overline{A_i \mathbf{x}})^{\mathrm{T}} (A_i \mathbf{x}) = \|A_i \mathbf{x}\|^2 \geq 0.$$

Here, the first equality follows from the definition of a Hermitian matrix.

Now we compute

$$
\begin{aligned}
0 = \bar{\mathbf{x}}^{\mathrm{T}} \mathcal{O} \mathbf{x} &= \bar{\mathbf{x}}^{\mathrm{T}} (A_1^2 + A_2^2 + \cdots + A_m^2) \mathbf{x} \\
&= \bar{\mathbf{x}}^{\mathrm{T}} A_1^2 \mathbf{x} + \bar{\mathbf{x}}^{\mathrm{T}} A_2^2 \mathbf{x} + \cdots + \bar{\mathbf{x}}^{\mathrm{T}} A_m^2 \mathbf{x} \\
&= \|A_1 \mathbf{x}\|^2 + \|A_2 \mathbf{x}\|^2 + \cdots + \|A_m \mathbf{x}\|^2.
\end{aligned}
$$

Since each length $\|A_i\mathbf{x}\|$ is a non-negative real number, this implies that we have $A_i\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathrm{R}^{\mathbf{n}}$ . Hence we must have $A_i = \mathcal{O}$ for each $i = 1, 2, \dots, m$.

---

13 Find the inverse matrix of the matrix

$$A = \begin{bmatrix} \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \\ -\frac{3}{7} & \frac{6}{7} & -\frac{2}{7} \end{bmatrix}.$$

## Hint.

You may use the augmented matrix method to find the inverse matrix.
Here we give an alternative way to find the inverse matrix by noting that $A$ is an orthogonal matrix.

Recall that a matrix $B$ is orthogonal if $B^{\mathrm{T}}B = B^{\mathrm{T}}B = I$.
Thus, once we know $B$ is an orthogonal matrix, then the inverse matrix $B^{-1}$ is just the transpose matrix $B^{\mathrm{T}}$.

Also, recall that a matrix $B$ is orthogonal if and only if the column vectors of $B$ form an orthonormal set.

## Solution.

We first show that $A$ is an orthogonal matrix.
To do this, it suffices to that the column vectors form an orthonormal set.

Let

$$\mathbf{v}_1 = \begin{bmatrix} \frac{2}{7} \\ \frac{6}{7} \\ -\frac{3}{7} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \frac{3}{7} \\ \frac{2}{7} \\ \frac{6}{7} \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} \frac{6}{7} \\ -\frac{3}{7} \\ -\frac{2}{7} \end{bmatrix}$$

be the column vectors of $A$.

Then the length of the vector $\mathbf{v}_1$ is

$$||\mathbf{v}_1|| = \sqrt{(2/7)^2 + (6/7)^2 + (-3/7)^2} = 1.$$

Similarly, we have $||\mathbf{v}_2|| = ||\mathbf{v}_3|| = 1$.
Thus, column vectors are unit vectors.

The dot (inner) product of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \frac{2}{7} \cdot \frac{3}{7} + \frac{6}{7} \cdot \frac{2}{7} + \left(-\frac{3}{7}\right) \cdot \frac{6}{7} = 0.$$

Similarly, we have

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = 0, \quad \mathbf{v}_2 \cdot \mathbf{v}_3 = 0.$$

Therefore, the column vectors are orthogonal.

Hence the column vectors of $A$ are orthonormal, and this implies that $A$ is an orthogonal matrix.

Namely, $A^{\mathrm{T}} = A^{-1}$.

Thus the inverse matrix of $A$ is

$$A^{-1} = \begin{bmatrix} \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \\ \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ \frac{6}{7} & -\frac{3}{7} & -\frac{2}{7} \end{bmatrix}.$$

---

14 A complex square $(n \times n)$ matrix $A$ is called **normal** if

$$A^* A = AA^*,$$

where $A^*$ denotes the conjugate transpose of $A$, that is $A^* = \bar{A}^{\mathrm{T}}$.

A matrix $A$ is said to be **nilpotent** if there exists a positive integer $k$ such that $A^k$ is the zero matrix.

**(a)** Prove that if $A$ is both normal and nilpotent, then $A$ is the zero matrix. You may use the fact that every normal matrix is diagonalizable.

**(b)** Give a proof of (a) without referring to eigenvalues and diagonalization.

**(c)** Let $A, B$ be $n \times n$ complex matrices. Prove that if $A$ is normal and $B$ is nilpotent such that $A + B = I$, then $A = I$, where $I$ is the $n \times n$ identity matrix.

## Proof.

### (a) If $A$ is normal and nilpotent, then $A = O$

Since $A$ is normal, it is diagonalizable. Thus there exists an invertible matrix $P$ such that $P^{-1}AP = D$, where $D$ is a diagonal matrix whose diagonal entries are eigenvalues of $A$.

Since $A$ is nilpotent, all the eigenvalues of $A$ are 0. (See the post "Nilpotent matrix and eigenvalues of the matrix" for the proof.)

Hence the diagonal entries of $D$ are zero, and we have $D = O$, the zero matrix.

It follows that we have

$$A = PDP^{-1} = POP^{-1} = O.$$

Therefore, every normal nilpotent matrix must be a zero matrix.

### (b) Give a proof of (a) without referring to eigenvalues and diagonalization.

Since $A$ is nilpotent, there exists a positive integer $k$ such that $A^k = O$.

We prove by induction on $k$ that $A = O$.

The base case $k = 1$ is trivial.

---

Suppose $k > 1$ and the case $k - 1$ holds. Let $B = A^{k-1}$. Note that since $A$ is normal, the matrix $B$ is also normal.

For any vector $x \in \mathrm{C}^n$, we compute the length of the vector $B^*Bx$ as follows.

$$\begin{aligned} \|B^*Bx\| &= (B^*Bx)^*(B^*Bx) && \text{by definition of the length} \\ &= x^*B^*(BB^*)Bx \\ &= x^*B^*(B^*B)Bx && \text{since } B \text{ is normal} \\ &= x^*(B^*)^2B^2x, \end{aligned}$$

and the last expression is $O$ since $B^2 = A^{2k-2} = O$ as $k \geq 2$ implies $2k - 2 \geq k$.

Hence we have $B^*Bx = \mathbf{0}$ for every $x \in \mathrm{C}^n$.

---

This yields that

$$\begin{aligned} \|Bx\| &= (Bx)^*(Bx) \\ &= x^*B^*Bx = 0, \end{aligned}$$

for every $x \in \mathrm{C}^n$, and hence $B = O$.

By the induction hypothesis, $A^{k-1} = O$ implies $A = O$, and the induction is completed.

So the matrix $A$ must be the zero matrix.

### (c) If $A$ is normal and $B$ is nilpotent such that $A + B = I$, then $A = I$

We claim that the matrix $B$ is normal as well. If this claim is proved, then part (a) yields that $B = O$ since $B$ is a nilpotent normal matrix, which implies the result $A = I$.

To prove the claim, we compute

$$\begin{aligned} B^*B &= (I - A)^*(I - A) \\ &= (I - A^*)(I - A) \\ &= I - A - A^* + A^*A, \end{aligned}$$

and

$$\begin{aligned} BB^* &= (I - A)(I - A)^* \\ &= (I - A)(I - A^*) \\ &= I - A^* - A + AA^* \\ &= I - A^* - A + A^*A && \text{since } A \text{ is normal.} \end{aligned}$$

It follows that we have $B^*B = BB^*$, and thus $B$ is normal.

Hence, the claim is proved.

# Eigenvalues and Eigenvectors of Linear Transformations

**Definition**

Let $T : V \to V$ be a linear transformation from a vector space $V$ to itself.

1. We say that $\lambda$ is an **eigenvalue** of $T$ if there exists a nonzero vector $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \lambda \mathbf{v}$.
2. For each eigenvalue $\lambda$ of $T$, nonzero vectors $\mathbf{v}$ satisfying $T(\mathbf{v}) = \lambda \mathbf{v}$ is called **eigenvectors** corresponding to $\lambda$.

## Problems

1. Let $T$ be the linear transformation from the vector space $\mathbb{R}^2$ to $\mathbb{R}^2$ itself given by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 + x_2 \\ x_1 + 3x_2 \end{bmatrix}.$$

**(a)** Verify that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

are eigenvectors of the linear transformation $T$, and conclude that $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $\mathbb{R}^2$ consisting of eigenvectors.

**(b)** Find the matrix of $T$ with respect to the basis $B = \{\mathbf{v}_1, \mathbf{v}_2\}$.

### Solution.

#### (a) $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $\mathbb{R}^2$ consisting of eigenvectors

We compute that

$$T(\mathbf{v}_1) = T\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2\mathbf{v}_1$$

and

$$T(\mathbf{v}_2) = T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\mathbf{v}_2.$$

Thus, $\mathbf{v}_1$ is an eigenvector corresponding to the eigenvalue $2$ and $\mathbf{v}_2$ is an eigenvector corresponding to the eigenvalue $4$.

Since $\mathbf{v}_1, \mathbf{v}_2$ are eigenvectors corresponding to distinct eigenvalues, they are linearly independent, and thus $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $\mathbb{R}^2$.

#### (b) Find the matrix of $T$ with respect to the basis $B$

From the computation in part (a), we have

$$T(\mathbf{v}_1) = 2\mathbf{v}_1 + 0\mathbf{v}_2$$
$$T(\mathbf{v}_2) = 0\mathbf{v}_1 + 4\mathbf{v}_2.$$

Hence the coordinate vectors of $T(\mathbf{v}_1), T(\mathbf{v}_2)$ with respect to the basis $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $\mathbb{R}^2$ are

$$[T(\mathbf{v}_1)]_B = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, [T(\mathbf{v}_2)]_B = \begin{bmatrix} 0 \\ 4 \end{bmatrix}.$$

Thus the matrix $A$ of the linear transformation $T$ with respect to the basis $B$ is

$$A = [\,[T(\mathbf{v}_1)]_B, [T(\mathbf{v}_2)]_B\,] = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

---

**2** Let $P_1$ be the vector space of all real polynomials of degree $1$ or less. Consider the linear transformation $T: P_1 \to P_1$ defined by $T(ax+b) = (3a+b)x + a + 3$ for any $ax+b \in P_1$.
**(a)** With respect to the basis $B = \{1, x\}$, find the matrix of the linear transformation $T$.
**(b)** Find a basis $B'$ of the vector space $P_1$ such that the matrix of $T$ with respect to $B'$ is a diagonal matrix.
**(c)** Express $f(x) = 5x + 3$ as a linear combination of basis vectors of $B'$.

---

(a) Find the matrix of the linear transformation $T$ with respect to the basis $B = \{1, x\}$ .

The matrix $A$ for $T$ with respect to the basis $B = \{1, x\}$ is given by

$$A = [\,[T(1)]_B, [T(x)]_B\,].$$

Since we have

$$T(1) = x + 3 \text{ and } T(x) = 3x + 1,$$

the coordinate vectors of these vector with respect to $B$ are

$$[T(1)]_B = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ and } [T(x)]_B = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Therefore, the matrix $A$ for $T$ with respect to $B$ is

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

(b) Find a basis $B'$ of the vector space $P_1$ such that the matrix of $T$ with respect to $B'$ is a diagonal matrix.

We first diagonalize the matrix $A$ obtained in (a).

Solving the characteristic polynomial

$$\begin{aligned} p_A(t) = \det(A - tI) &= \begin{bmatrix} 3-t & 1 \\ 1 & 3-t \end{bmatrix} \\ &= t^2 - 6t + 8 = (t-2)(t-4), \end{aligned}$$

the eigenvalues are $2$ and $4$. Since $A$ has two distinct eigenvalues, it is diagonalizable.

We find the eigenvectors corresponding to the eigenvalue $2$.
The eigenvectors are nonzero solutions of $(A-2I)\mathrm{x}=0$ .
From this, we find that $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector.
Similarly, we find that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $4$.

Thus, if we set

$$P=\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

then we have the diagonalization of $A$ :

$$\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}=P^{-1}AP.$$

It follows that $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ is the matrix representation for $T$ with respect to the basis

$$B'=\{\mathrm{v}_1,\mathrm{v}_2\},$$

where

$$[\mathrm{v}_1]_B=\begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } [\mathrm{v}_2]_B=\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and $P$ is the transition matrix (change of basis) from $B'$ to $B$.

(Note that

$$[\mathrm{v}_1]_B=P[\mathrm{v}_1]_{B'}=P\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Similarly for $[\mathrm{v}_2]_B$ .)

Hence, $\mathrm{v}_1=-x+1$ and $\mathrm{v}_2=x+1$ form the basis $B'=\{\mathrm{v}_1,\mathrm{v}_2\}$ with respect to which the matrix for $T$ is the diagonal matrix $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$.

> (c) Express $f(x)=5x+3$ as a linear combination of basis vectors of $B'$.

Note that since $P$ is the transition matrix from $B'$ to $B$, we have the following relation between coordinate vectors. For any $\mathrm{v}\in P_1$, we have

$$P[\mathrm{v}]_{B'}=[\mathrm{v}]_B.$$

Using this, we have

$$\begin{aligned}[f(x)]_{B'}&=P^{-1}[f(x)]_B\\ &=\frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix}=\begin{bmatrix} -1 \\ 4 \end{bmatrix}.\end{aligned}$$

Hence we obtain the linear combination of $f(x)$ with $\mathrm{v}_1, \mathrm{v}_2$:

$$f(x) = (-1)\mathrm{v}_1 + 4\mathrm{v}_2$$

(You may confirm that $5x + 3 = -(-x+1) + 4(x+1)$ .)

③ Let $V$ be a real vector space of all real sequences $(a_i)_{i=1}^{\infty} = (a_1, a_2, \ldots)$. Let $U$ be a subspace of $V$ defined by

$$U = \{(a_i)_{i=1}^{\infty} \in V \mid a_{n+2} = 2a_{n+1} + 3a_n \text{ for } n = 1, 2, \ldots\}.$$

Let $T$ be the linear transformation from $U$ to $U$ defined by

$$T\big((a_1, a_2, \ldots)\big) = (a_2, a_3, \ldots).$$

**(a)** Find the eigenvalues and eigenvectors of the linear transformation $T$.
**(b)** Use the result of (a), find a sequence $(a_i)_{i=1}^{\infty}$ satisfying $a_1 = 2, a_2 = 7$.

---

### (a) Find the eigenvalues and eigenvectors of the linear transformation $T$.

Note that each sequence $(a_i)_{i=1}^{\infty}$ in $U$ is determined by the first two numbers $a_1, a_2$ since the rests can be obtained from the linear recurrence relation $a_{n+2} = 2a_{n+1} + 3a_n$ .
Thus, for example, we take $(a_1, a_2) = (1, 0), (0, 1)$ and obtain a basis $B = \{\mathrm{u}_1, \mathrm{u}_2\}$ of $U$, where

$$\begin{aligned} \mathrm{u}_1 &= (1, 0, 3, 6, 21, \cdots) \\ \mathrm{u}_2 &= (0, 1, 2, 7, 20, \ldots). \end{aligned}$$

We find the matrix $A$ of $T$ with respect to the basis $B$ .

We have

$$\begin{aligned} T(\mathrm{u}_1) &= (0, 3, 6, 21, \cdots) = 0\mathrm{u}_1 + 3\mathrm{u}_2 \\ T(\mathrm{u}_2) &= (1, 2, 7, 20, \ldots) = \mathrm{u}_1 + 2\mathrm{u}_2. \end{aligned}$$

Thus, the matrix $A$ for $T$ is

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}.$$

Then the eigenvalues of $T$ are eigenvalues of $A$ .

The characteristic polynomial for $A$ is

$$p_A(t) = \det(A - tI) = t^2 - 2t - 3 = (t+1)(t-3).$$

Thus the eigenvalues are $t = -1, 3$ .

Let us find eigenvectors corresponding to $t = -1$ .

We solve

$$T\big((a_1, a_2, \ldots)\big) = -(a_1, a_2, \ldots),$$

or equivalently,

$$(a_2, a_3, \ldots) = -(a_1, a_2, \ldots).$$

This yields the relation

$$a_{n+1} = -a_n$$

for $n = 1, 2, \dots$.

Hence the sequence $(a_i)_{i=1}^{\infty}$ is a geometric progression with initial value $a_1$ and ratio $-1$.
Therefore, the sequence $(a_i)_{i=1}^{\infty}$ such that

$$a_i = (-1)^{i-1} a_1$$

are eigenvectors corresponding to the eigenvalue $-1$.

Similarly, eigenvectors corresponding to the eigenvalue $3$ are geometric progression with initial value $a_1$ and ratio $3$, thus

$$a_i = 3^{i-1} a_1.$$

(b) Find a sequence $(a_i)_{i=1}^{\infty}$ satisfying $a_1 = 2, a_2 = 7$.

From part (a), we see that sequences

$$\mathrm{v}_1 = \left((-1)^{i-1}\right)_{i=1}^{\infty} \text{ and } \mathrm{v}_2 = \left(3^{i-1}\right)_{i=1}^{\infty}$$

form a basis of $U$ consisting of eigenvectors of $T$. (We chose $a_1 = 1$.)
Hence any sequence $(a_i)_{i=1}^{\infty}$ can be written as a linear combination of these two vectors:

$$\begin{aligned}(a_i)_{i=1}^{\infty} &= c_1 \mathrm{v}_1 + c_2 \mathrm{v}_2 \\ &= c_1 \left((-1)^{i-1}\right)_{i=1}^{\infty} + c_2 \left(3^{i-1}\right)_{i=1}^{\infty}\end{aligned}$$

for some $c_1, c_2$.

If the sequence satisfies the initial condition $a_1 = 2, a_2 = 7$, then we have

$$\begin{aligned}2 &= a_1 = c_1 + c_2 \\ 7 &= a_2 = -c_1 + 3c_2.\end{aligned}$$

Solving this, we obtain $c_1 = -1/4$ and $c_2 = 9/4$.
Hence we have

$$\begin{aligned}(a_i)_{i=1}^{\infty} &= -\frac{1}{4}\left((-1)^{i-1}\right)_{i=1}^{\infty} + \frac{9}{4}\left(3^{i-1}\right)_{i=1}^{\infty} \\ &= \frac{1}{4}\left((-1)^{i}\right)_{i=1}^{\infty} + \frac{1}{4}\left(3^{i+1}\right)_{i=1}^{\infty} \\ &= \frac{1}{4}\left((-1)^{i} + 3^{i+1}\right)_{i=1}^{\infty}.\end{aligned}$$

In summary, the sequence $(a_i)_{i=1}^{\infty}$ satisfying the linear recurrence relation $a_{n+2} = 2a_{n+1} + 3a_n$ with initial conditions $a_1 = 2, a_2 = 7$ is

$$(a_i)_{i=1}^{\infty} = \frac{1}{4}\left((-1)^i + 3^{i+1}\right)_{i=1}^{\infty}.$$

Check out the following problems about how to solve a linear recurrence relation using the vector space techniques, like the current problem.

④ Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation and let $A$ be the matrix representation of $T$ with respect to the standard basis of $\mathbb{R}^2$. Prove that the following two statements are equivalent.
**(a)** There are exactly two distinct lines $L_1, L_2$ in $\mathbb{R}^2$ passing through the origin that are mapped onto themselves: $T(L_1) = L_1$ and $T(L_2) = L_2$.
**(b)** The matrix $A$ has two distinct nonzero real eigenvalues.

## Proof.

### (a) $\implies$ (b).

Suppose that the statement (a) holds.
That is, there are lines $L_1, L_2$ in $\mathbb{R}^2$ passing through the origin that are mapped onto themselves, and no other lines passing through the origin are mapped onto themselves.

A line passing through the origin in $\mathbb{R}^2$ is a one-dimensional subspace in $\mathbb{R}^2$. Thus, it is spanned by a single nonzero vector.

We have

$$L_1 = \operatorname{Span}(\mathbf{v}_1) \text{ and } L_2 = \operatorname{Span}(\mathbf{v}_2)$$

for some nonzero vectors $\mathbf{v}_1, \mathbf{v}_2$.

Since $T(L_1) = L_1$, we have $A\mathbf{v}_1 \in L_1 = \operatorname{Span}(\mathbf{v}_1)$.
Hence there exists $\lambda_1 \in \mathbb{R}$ such that

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1.$$

Since $\mathbf{v}_1$ is a nonzero vector, this means that $\lambda_1$ is an eigenvalue of $A$.

We claim that $\lambda_1 \neq 0$. Otherwise, we have $A\mathbf{v}_1 = \mathbf{0}$.
Any vector in $L_1$ is of the form $\mathbf{v} = a\mathbf{v}_1$ for some $a \in \mathbb{R}$. So we have

$$A\mathbf{v} = A(a\mathbf{v}_1) = aA\mathbf{v}_1 = \mathbf{0},$$

and this yields that $T(L_1) = \{\mathbf{0}\}$, a contradiction.
Hence $\lambda_1$ is a nonzero real eigenvalue of $A$.

By the same argument, there is a nonzero real eigenvalue $\lambda_2$ such that $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$.

It remains to show that $\lambda_1 \neq \lambda_2$.
Assume on the contrary that $\lambda := \lambda_1 = \lambda_2$.

Since $\mathbf{v}_1, \mathbf{v}_2$ are a basis of two distinct lines, they form a basis of $\mathbb{R}^2$.
Hence any vector $\mathbf{v} \in \mathbb{R}^2$ can be written as a linear combination

$$\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$$

for some $a, b \in \mathbb{R}$.

Then we have

$$\begin{aligned} A\mathbf{v} &= A(a\mathbf{v}_1 + b\mathbf{v}_2) \\ &= aA\mathbf{v}_1 + bA\mathbf{v}_2 \\ &= a\lambda\mathbf{v}_1 + b\lambda\mathbf{v}_2 \\ &= \lambda(a\mathbf{v}_1 + b\mathbf{v}_2) \\ &= \lambda\mathbf{v}. \end{aligned}$$

---

This implies that any line spanned by a nonzero vector $\mathbf{v}$ is mapped onto itself by the linear transformation $T$.
This contradicts our assumption that $L_1, L_2$ are the only such lines.

Therefore, $\lambda_1 \neq \lambda_2$, and we have proved that $A$ has two distinct nonzero real eigenvalues.

> (b) $\implies$ (a).

Now we suppose that the statement (b) holds.
Namely, we suppose that there are two distinct nonzero real eigenvalues of $A$. Let us call them $\lambda_1, \lambda_2$ and let $\mathbf{v}_1, \mathbf{v}_2$ be eigenvectors corresponding to $\lambda_1, \lambda_2$, respectively.

---

In general eigenvectors corresponding to distinct eigenvalues are linearly independent.
Thus, $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent.
Hence the lines $L_1, L_2$ spanned by $\mathbf{v}_1, \mathbf{v}_2$ are distinct.

Each vector on $L_1$ is of the form $\mathbf{v} = a\mathbf{v}_1$ for some $a \in R$. Hence we have

$$A\mathbf{v} = A(a\mathbf{v}_1) = aA\mathbf{v}_1 = a\lambda_1\mathbf{v}_1 = \lambda\mathbf{v}.$$

It follows that each vector $\mathbf{v} \in L_1$ is mapped onto $L_1$.
(Since $\lambda \neq 0$, the image under $T$ is one-dimensional, hence $T(L_1) = L_1$.)
Similarly, we have $T(L_2) = L_2$.

---

Finally, we show that no other lines passing through the origin are mapped onto themselves by $T$.

Assume that we have a line $L = \operatorname{Span}(\mathbf{w})$ such that $T(L) = L$.
Then since we have $A\mathbf{w} \in L = \operatorname{Span}(\mathbf{w})$, there is $\mu \in \mathbb{R}$ such that

$$A\mathbf{w} = \mu\mathbf{w}. \tag{*}$$

Since $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $\mathbb{R}^2$, we have

$$\mathbf{w} = a\mathbf{v}_1 + b\mathbf{v}_2$$

for some $a, b \in \mathbb{R}$.

Then the left hand side of (*) becomes

$$\begin{aligned} A\mathrm{w} &= A(a\mathrm{v}_1 + b\mathrm{v}_2) \\ &= aA\mathrm{v}_1 + bA\mathrm{v}_2 \\ &= a\lambda_1\mathrm{v}_1 + b\lambda_2\mathrm{v}_2. \end{aligned}$$

---

Thus the relation (*) yields that

$$\begin{aligned} a\lambda_1\mathrm{v}_1 + b\lambda_2\mathrm{v}_2 = \mu a\mathrm{v}_1 + \mu b\mathrm{v}_2 \\ a(\lambda_1 - \mu)\mathrm{v}_1 + b(\lambda_2 - \mu)\mathrm{v}_2 = 0. \end{aligned}$$

Since $\mathrm{v}_1, \mathrm{v}_2$ are linearly independent, we have

$$a(\lambda_1 - \mu) = 0 \text{ and } b(\lambda_2 - \mu) = 0.$$

The equality $a(\lambda_1 - \mu) = 0$ implies that either $a = 0$ or $\lambda_1 = \mu$.

If $a = 0$, then we have $\mathrm{w} = b\mathrm{v}_2$, and hence $L = L_2$.

If $\mu = \lambda_1$, then the second equality $b(\lambda_2 - \mu) = 0$ implies that $b = 0$ since $\lambda_1 \neq \lambda_2$.

Then we have $\mathrm{w} = a\mathrm{v}_1$, and hence $L = L_1$.

In either case, $L$ must be $L_1$ or $L_2$.

This completes the proof.

⑤ Let $V$ be a real vector space of all real sequences $(a_i)_{i=1}^{\infty} = (a_1, a_2, \dots)$.
Let $U$ be the subspace of $V$ consisting of all real sequences that satisfy the linear recurrence relation $a_{k+2} - 5a_{k+1} + 3a_k = 0$ for $k = 1, 2, \dots$. Let $T$ be the linear transformation from $U$ to $U$ defined by

$$T\big((a_1, a_2, \dots)\big) = (a_2, a_3, \dots).$$

Let $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ be a basis of $U$, where

$$\begin{aligned} \mathbf{u}_1 &= (1, 0, -3, -15, -66, \dots) \\ \mathbf{u}_2 &= (0, 1, 5, 22, 95, \dots). \end{aligned}$$

Let $A$ be the matrix representation of the linear transformation $T : U \to U$ with respect to the basis $B$.

**(a)** Find the eigenvalues and eigenvectors of $T$.

**(b)** Use the result of (a), find a sequence $(a_i)_{i=1}^{\infty}$ satisfying the linear recurrence relation $a_{k+2} - 5a_{k+1} + 3a_k = 0$ and the initial condition $a_1 = 1, a_2 = 1$.

**(c)** Find the formula for the sequences $(a_i)_{i=1}^{\infty}$ satisfying the linear recurrence relation $a_{k+2} - 5a_{k+1} + 3a_k = 0$ and express it using $a_1, a_2$.

## Solution.

### (a) Find the eigenvalues and eigenvectors of the linear transformation $T$.

Note that each sequence $(a_i)_{i=1}^{\infty}$ in $U$ is determined by the first two numbers $a_1, a_2$ since the rests can be obtained from the linear recurrence relation $a_{n+2} = 2a_{n+1} + 3a_n$.

Thus, for example, we take $(a_1, a_2) = (1, 0), (0, 1)$ and obtain a basis $B = \{\mathrm{u}_1, \mathrm{u}_2\}$ of $U$, where

$$\begin{aligned} \mathbf{u}_1 &= (1, 0, 3, 6, 21, \cdots) \\ \mathbf{u}_2 &= (0, 1, 2, 7, 20, \dots). \end{aligned}$$

We find the matrix $A$ of $T$ with respect to the basis $B$.

We have

$$\begin{aligned} T(\mathbf{u}_1) &= (0, 3, 6, 21, \cdots) = 0\mathbf{u}_1 + 3\mathbf{u}_2 \\ T(\mathbf{u}_2) &= (1, 2, 7, 20, \dots) = \mathbf{u}_1 + 2\mathbf{u}_2. \end{aligned}$$

Thus, the matrix $A$ for $T$ is

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}.$$

Then the eigenvalues of $T$ are eigenvalues of $A$.

The characteristic polynomial for $A$ is

$$p_A(t) = \det(A - tI) = t^2 - 2t - 3 = (t + 1)(t - 3).$$

Thus the eigenvalues are $t = -1, 3$.

Let us find eigenvectors corresponding to $t = -1$.

We solve

$$T\big((a_1, a_2, \dots)\big) = -(a_1, a_2, \dots),$$

or equivalently,

$$(a_2, a_3, \dots) = -(a_1, a_2, \dots).$$

This yields the relation

$$a_{n+1} = -a_n$$

for $n = 1, 2, \dots$.

Hence the sequence $(a_i)_{i=1}^{\infty}$ is a geometric progression with initial value $a_1$ and ratio $-1$.

Therefore, the sequence $(a_i)_{i=1}^{\infty}$ such that

$$a_i = (-1)^{i-1} a_1$$

are eigenvectors corresponding to the eigenvalue $-1$.

Similarly, eigenvectors corresponding to the eigenvalue $3$ are geometric progression with initial value $a_1$ and ratio $3$, thus

$$a_i = 3^{i-1} a_1.$$

(b) Find a sequence $(a_i)_{i=1}^{\infty}$ satisfying $a_1 = 2, a_2 = 7$.

From part (a), we see that sequences

$$\mathbf{v}_1 = \left((-1)^{i-1}\right)_{i=1}^{\infty} \text{ and } \mathbf{v}_2 = \left(3^{i-1}\right)_{i=1}^{\infty}$$

form a basis of $U$ consisting of eigenvectors of $T$. (We chose $a_1 = 1$.)
Hence any sequence $(a_i)_{i=1}^{\infty}$ can be written as a linear combination of these two vectors:

$$\begin{aligned}(a_i)_{i=1}^{\infty} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \\ &= c_1\left((-1)^{i-1}\right)_{i=1}^{\infty} + c_2\left(3^{i-1}\right)_{i=1}^{\infty}\end{aligned}$$

for some $c_1, c_2$.

If the sequence satisfies the initial condition $a_1 = 2, a_2 = 7$, then we have

$$\begin{aligned}2 &= a_1 = c_1 + c_2 \\ 7 &= a_2 = -c_1 + 3c_2.\end{aligned}$$

Solving this, we obtain $c_1 = -1/4$ and $c_2 = 9/4$.
Hence we have

$$\begin{aligned}(a_i)_{i=1}^{\infty} &= -\frac{1}{4}\left((-1)^{i-1}\right)_{i=1}^{\infty} + \frac{9}{4}\left(3^{i-1}\right)_{i=1}^{\infty} \\ &= \frac{1}{4}\left((-1)^{i}\right)_{i=1}^{\infty} + \frac{1}{4}\left(3^{i+1}\right)_{i=1}^{\infty} \\ &= \frac{1}{4}\left((-1)^{i} + 3^{i+1}\right)_{i=1}^{\infty}.\end{aligned}$$

In summary, the sequence $(a_i)_{i=1}^{\infty}$ satisfying the linear recurrence relation $a_{n+2} = 2a_{n+1} + 3a_n$ with initial conditions $a_1 = 2, a_2 = 7$ is

$$(a_i)_{i=1}^{\infty} = \frac{1}{4}\left((-1)^{i} + 3^{i+1}\right)_{i=1}^{\infty}.$$

# Jordan Canonical Form

## Problems

① Let $A$ be an $n \times n$ matrix such that $\operatorname{tr}(A^n) = 0$ for all $n \in \mathbb{N}$. Then prove that $A$ is a nilpotent matrix. Namely there exist a positive integer $m$ such that $A^m$ is the zero matrix.

### Steps.

② Use the Jordan canonical form of the matrix $A$.

③ We want to show that all eigenvalues are zero. (Review Nilpotent matrix and eigenvalues of the matrix)

④ Seeking a contradiction, assume some eigenvalues are not zero.

⑤ Using $\operatorname{tr}(A^n) = 0$, create a system of linear equations.

⑥ Calculate the determinant of the coefficient matrix of the system using the Vandermonde matrix formula.

⑦ Find a contradiction.

### Proof.

We first want to prove that all the eigenvalues of $A$ must be zero.
Seeking a contradiction, assume that some of the eigenvalues of $A$ are not zero.
So assume that $\lambda_i, i = 1, \dots, r$ are distinct nonzero eigenvalues of $A$ and each $m_i \geq 1$ is a multiplicity of $\lambda_i$.

We use the Jordan canonical form of the matrix $A$.
There exists an invertible matrix $S$ such that $S^{-1}AS = T$, where $T$ is an upper triangular matrix.The diagonal entries are eigenvalues of $A$.

Then we have for any positive integer $n$,

$$\begin{aligned}0 &= \mathrm{tr}(A^n) = \mathrm{tr}((STS^{-1})^n) = \mathrm{tr}(ST^nS^{-1}) = \mathrm{tr}(T^n)\\ &= m_1\lambda_1^n + m_2\lambda_2^n + \cdots + m_r\lambda_r^n.\end{aligned}$$

Note that since $T$ is an upper triangular matrix, the nonzero diagonal entries of $T^n$ are $\lambda_i^n$ appearing $m_i$ times.

Changing $n$ from $1$ to $r$, we obtain the system of linear equation. (Think $m_i$ as variables.)

$$\begin{aligned}m_1\lambda_1^1 + m_2\lambda_2^1 + \cdots + m_r\lambda_r^1 &= 0\\ m_1\lambda_1^2 + m_2\lambda_2^2 + \cdots + m_r\lambda_r^2 &= 0\\ &\vdots\\ m_1\lambda_1^r + m_2\lambda_2^r + \cdots + m_r\lambda_r^r &= 0\end{aligned}$$

Equivalently, we have the matrix equation

$$\begin{bmatrix} \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_r^1 \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_r^2 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^r & \lambda_2^r & \cdots & \lambda_r^r \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{*}$$

Let $B$ denote the matrix above whose entries are powers of eigenvalues $\lambda_i$.

We calculate the determinant of $B$.

$$\begin{aligned}\det(B) &= \lambda_1\lambda_2\cdots\lambda_r \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \cdots & \lambda_r^{r-1} \end{bmatrix}\\ &= \lambda_1\lambda_2\cdots\lambda_r \prod_{1\leq i<j\leq n} (\lambda_j - \lambda_i) \neq 0.\end{aligned}$$

(Note that the matrix above is a Vandermonde matrix.)

Thus the matrix $B$ is invertible. This means that the matrix equation (*) has unique solution

$$\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{*}$$

But this is a contradiction because each multiplicity $m_i$ is grater than zero.

Thus we proved that all eigenvalues of $A$ are zero.

Recall that a matrix is nilpotent if and only if its eigenvalues are zero.

See the post ↴

Nilpotent Matrix and Eigenvalues of the Matrix

for a proof of this fact.

Hence it follows from this fact that $A$ is a nilpotent matrix.

8 Let $A$ be $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be all the eigenvalues of $A$. (Some of them may be the same.) For each positive integer $k$, prove that $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ are all the eigenvalues of $A^k$.

By the triangularization (or Jordan canonical form), there exists a nonsingular matrix $S$ such that

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & \lambda_2 & * & * & * \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & \lambda_{n-1} & * \\ 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix}.$$

Here the right matrix is an upper triangular matrix whose diagonal entries are eigenvalues of $A$.

Then we have

$$S^{-1}A^kS = (S^{-1}AS)^k = \begin{bmatrix} \lambda_1^k & * & * & * & * \\ 0 & \lambda_2^k & * & * & * \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & \lambda_{n-1}^k & * \\ 0 & 0 & 0 & 0 & \lambda_n^k \end{bmatrix}.$$

---

The characteristic polynomial of the matrix $A^k$ is given by

$$\begin{aligned} p(t) &= \det(A^k - tI) \\ &= \det(S^{-1})\det(A^k - tI)\det(S) \\ &= \det(S^{-1}(A^k - tI)S) \\ &= \det(S^{-1}A^kS - tI) \\ &= \begin{vmatrix} \lambda_1^k - t & * & * & * & * \\ 0 & \lambda_2^k - t & * & * & * \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & \lambda_{n-1}^k - t & * \\ 0 & 0 & 0 & 0 & \lambda_n^k - t \end{vmatrix} \\ &= \prod_{i=1}^{n} (\lambda_i^k - t). \end{aligned}$$

Since the roots of the characteristic polynomial are all the eigenvalues, we see that $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ are all the eigenvalues of $A^k$.

9 Let $A$ be an $n \times n$ matrix and suppose that $A^r = I_n$ for some positive integer $r$. Then show that:

**(a)** $|\mathrm{tr}(A)| \leq n$.

**(b)** If $|\mathrm{tr}(A)| = n$, then $A = \zeta I_n$ for an $r$-th root of unity $\zeta$.

**(c)** $\mathrm{tr}(A) = n$ if and only if $A = I_n$.

Proof.

### (a) $|\mathrm{tr}(A)| \leq n$ .

We fist show that the eigenvalues of $A$ are $r$-th roots of unity.

Let $\lambda$ be an eigenvalue of $A$ and let $\mathbf{x}$ be an eigenvector corresponding to $\lambda$. That is, $A\mathbf{x} = \lambda\mathbf{x}$ .

Multiplying this equality by $A$ on the left, we see that

$$A^2\mathbf{x} = \lambda A\mathbf{x} = \lambda^2\mathbf{x}.$$

Again multiplying this equality by $A$ on the left we get $A^3\mathbf{x} = \lambda^3\mathbf{x}$ .

Inductively, we obtain $A^r\mathbf{x} = \lambda^r\mathbf{x}$ .
Since $A^r = I_n$ , we get $\mathbf{x} = \lambda^r\mathbf{x}$ , Thus $(\lambda^r - 1)\mathbf{x} = 0$ and since $\mathbf{x}$ is nonzero vector (as it is an eigenvector), we have $\lambda^r = 1$. Therefore eigenvalues are $r$-th roots of unity.

---

Next, we consider the Jordan canonical form of $A$ . There exists an invertible matrix $P$ such that $P^{-1}AP$ is an upper triangular matrix $T = (t_{ij})$ whose diagonal entries are eigenvalues of $A$ .
Since the trace of $A$ is equal to the trace of $P^{-1}AP$ , we see that $\mathrm{tr}(A)$ is a sum of $n$ $r$-th roots of unity. Thus

$$|\mathrm{tr}(A)| = |\sum_{i=1}^{n} t_{ii}| \leq \sum_{i=1}^{n} |t_{ii}| = n. \qquad (*)$$

### (b) If $|\mathrm{tr}(A)| = n$ , then $A = \zeta I_n$ for an $r$-th root of unity $\zeta$

We first refine the proof of part (a).
Note that the matrix $A$ is a solution of the equation $x^r - 1$,

So the minimal polynomial of $A$ divides $x^r - 1$, thus the minimal polynomial has no repeated roots.
Thereby the Jordan canonical form $T$ is actually a diagonal matrix.

Now we start the proof of (b). The inequality in (*) becomes the equality if and only if all the roots of unity $t_{ii}$ are the same root of unity $\zeta$.
Then the Jordan canonical form is $T = \zeta I_n$ , hence $A = \zeta I_n$ .

### (c) $\mathrm{tr}(A) = n$ if and only if $A = I_n$

By part (b), we have $A = \zeta I_n$ for some $r$-th root of unity $\zeta$ but then $n = \mathrm{tr}(A) = \zeta n$ implies $\zeta = 1$.
Thus $A = I_n$ as required.

⑩ Let $A$ be an $n \times n$ matrix such that $A^k = I_n$, where $k \in \mathbb{N}$ and $I_n$ is the $n \times n$ identity matrix. Show that the trace of $(A^{-1})^{\mathrm{T}}$ is the conjugate of the trace of $A$. That is, show that $\mathrm{tr}((A^{-1})^{\mathrm{T}}) = \overline{\mathrm{tr}(A)}$.

### Hint.

⑪ Note that $\mathrm{tr}(B) = \mathrm{tr}(B^{\mathrm{T}})$ for any square matrix $B$ .

⑫ Use the Jordan canonical form of $A$ .

⑬ Show that eigenvalues are $k$-th roots of unity.

### Proof.

First note that the trace of a matrix is the same as the trace of its transpose. Thus we only have to show $\operatorname{tr}(A^{-1}) = \overline{\operatorname{tr}(A)}$ .

There is an invertible matrix $P$ such that $P^{-1}AP$ is the Jordan canonical form. That is, $P^{-1}AP = T$ , where $T$ is an upper triangular matrix whose diagonal entries are eigenvalues of $A$ .

Since $A$ is invertible (to see this take the determinant of $A^k = I_n$ ), the matrix $T$ is also invertible and $P^{-1}A^{-1}P = T^{-1}$. Then we have

$$\begin{aligned} \operatorname{tr}(A) &= \operatorname{tr}(P^{-1}AP) = \operatorname{tr}(T) \\ \operatorname{tr}(A^{-1}) &= \operatorname{tr}(P^{-1}A^{-1}P) = \operatorname{tr}(T^{-1}) \end{aligned}$$

Now let $\lambda_1, \lambda_2, \dots, \lambda_n$ be eigenvalues of $A$ . Then the upper triangular matrix $T$ and its inverse matrix are

$$T = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}, \; T^{-1} = \begin{bmatrix} \lambda_1^{-1} & * & * & * \\ 0 & \lambda_2^{-1} & * & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^{-1} \end{bmatrix}$$

Thus we have

$$\begin{aligned} \operatorname{tr}(A) &= \lambda_1 + \lambda_2 + \cdots + \lambda_n \\ \operatorname{tr}(A^{-1}) &= \lambda_1^{-1} + \lambda_2^{-1} + \cdots + \lambda_n^{-1} \end{aligned}$$

Next, we show that $\lambda_i^{-1} = \overline{\lambda_i}$ for $i = 1, 2, \dots, n$ .
This follows from the fact that if $A^k = I_n$ then the eigenvalues are $k$-th roots of unity.

Assuming this, we have $1 = |\lambda_i| = \overline{\lambda_i}\lambda_i$ , hence $\lambda_i^{-1} = \overline{\lambda_i}$ .

To prove the fact, let $\lambda$ be an eigenvalue of $A$ and let $x$ be an eigenvector corresponding to $\lambda$ .

Then we have $Ax = \lambda x$ . Using this relation successively we have

$$x = I_n x = A^k x = \lambda A^{k-1} x = \lambda^2 A^{k-2} x = \cdots = \lambda^k x$$

Since $x$ is a nonzero vector, we have $\lambda^k = 1$, and $\lambda$ is a $k$-th root of unity.

Now we have

$$\begin{aligned}\operatorname{tr}(A^{-1}) &= \lambda_1^{-1}+\lambda_2^{-1}+\cdots+\lambda_n^{-1}\\ &= \overline{\lambda_1}+\overline{\lambda_2}+\cdots+\overline{\lambda_n}\\ &= \overline{\lambda_1+\lambda_2+\cdots+\lambda_n}\\ &= \overline{\operatorname{tr}(A)}\,.\end{aligned}$$

This completes the proof.

⑭ **(a)** Does there exist a $2 \times 2$ matrix $A$ with $A^3 = O$ but $A^2 \neq O$? Here $O$ denotes the $2 \times 2$ zero matrix.
**(b)** Does there exist a $3 \times 3$ real matrix $B$ such that $B^2 = A$ where

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \ ?$$

(*Princeton University*)

## Hint.

⑮ (a) Consider the eigenvalues and the Jordan canonical form of $A$

⑯ (b) Diagonalize the matrix $A$

## Solution.

### (a) Does there exist a $2 \times 2$ matrix $A$ with $A^3 = O$ but $A^2 \neq O$?

We claim that there is no such $2 \times 2$ matrix $A$.

Suppose that we have $A^3 = O$. This implies that all the eigenvalues of $A$ are zero.

To see this, let $\lambda$ be an eigenvalue of $A$ and let $\mathbf{x}$ be a corresponding eigenvector. Then applying the defining relation $A\mathbf{x} = \lambda\mathbf{x}$ successively, we have

$$0 = A^3\mathbf{x} = \lambda A^2\mathbf{x} = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}.$$

Thus, the eigenvalue $\lambda$ must be zero since $\mathbf{x}$ is a nonzero vector since it is an eigenvector.

Therefore, there exists an invertible matrix $P$ such that the Jordan canonical form of $A$ is

$P^{-1}AP = \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}$ for some number $a$.

Then $A = P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}P^{-1}$ and we have

$$\begin{aligned}A^2 &= \left(P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}P^{-1}\right)\left(P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}P^{-1}\right) = P\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}^2 P^{-1}\\ &= P\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}P^{-1} = O.\end{aligned}$$

Hence $A^2$ must be zero, and we conclude that there is no $2 \times 2$ matrix $A$ satisfying both $A^3 = O$ and $A^2 \neq O$.

(b) Does there exist a $3 \times 3$ real matrix $B$ such that $B^2 = A$?

The answer is yes and we will construct such a matrix $B$ as follows.

The characteristic polynomial $p(t)$ of the matrix $A$ is

$$p(t) := \det(A - tI) = -t(t-1)(t-3).$$

Thus the eigenvalues are $\lambda = 0, 1, 3$.
The corresponding eigenvalues are obtained by solving the system $(A - \lambda I)\mathbf{x} = 0$ and we see that

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

are eigenvectors corresponding to eigenvalues $0, 1, 3$, respectively.

Thus the matrix

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix}$$

is invertible and diagonalize the matrix $A$ such that

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Then if we have the equality $B^2 = A$, then we need to have

$$(P^{-1}BP)(P^{-1}BP) = p^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

From this we see that the matrix

$$B = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} P^{-1}$$

satisfies the above relation.

In fact, we compute

$$B^2 = \left( P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} P^{-1} \right) \left( P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} P^{-1} \right)$$

$$= P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}^2 P^{-1} = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} P^{-1} = A.$$

Remark that the matrix $B$ is real since $P$ is real. Thus the matrix $B$ satisfies the conditions of the problem.

## Remark.

The explicit form of the matrix $P$ was used to assure that the matrix $B$ is real.
So we did not have to compute the matrix $P$ explicitly if we use the fact that the symmetric matrix is diagonalizable by a real orthogonal matrix.

## Generalization.

See problem A square root matrix of a symmetric matrix with non-negative eigenvalues for a more general question than part (b).

Furthermore, try the next problem.

**Problem**.
Prove that a positive definite matrix has a unique positive definite square root.

For a solution of this problem, see the post
A Positive Definite Matrix Has a Unique Positive Definite Square Root